REPORT DOCUMENTATION PAGE

Form Approved
OMB NO. 0704-0188

1. AGENCY USE ONLY (Leave Blank)	2. REPORT DATE 28 JUNE 02	3. REPORT TYPE AND DATES COVERED Final Progress Report, 1999-2002

4. TITLE AND SUBTITLE
Information-Theoretic Information Fusion: Final Progress Report

5. FUNDING NUMBERS
DAAG55-98-C-0039

6. AUTHOR(S)
Ronald P.S. Mahler

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Lockheed Martin Tactical Systems, 3333 Pilot Knob Road, Eagan MN 55121

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U. S. Army Research Office
P.O. Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

37629-EL
.17

11. SUPPLEMENTARY NOTES
The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision, unless so designated by other documentation.

12 a. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for public release; distribution unlimited.

12 b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

This Final Report summarizes research on information fusion based on finite-set statistics (FISST). FISST provides a fully unified, scientifically defensible, probabilistic foundation for the following aspects of multisource, multitarget, multiplatform data fusion: (1) multisource integration (detection, identification, and tracking) based on Bayesian filtering and estimation; (2) sensor management using control theory; (3) performance evaluation using information theory; (4) expert-systems theory (fuzzy logic, the Dempster-Shafer theory of evidence, rule-based inference); (5) distributed fusion; and (5) aspects of situation/ threat assessment. The core of FISST is a multisource-multitarget differential and integral calculus based on the fact that belief-mass functions are the multisensor-multitarget counterparts of probability-mass functions. One purpose of this calculus is to enable signal processing engineers to directly generalize conventional, engineering-friendly statistical reasoning to multisensor, multitarget, multi-evidence applications. A second purpose is to extend Bayesian (and other probabilistic) methodologies so that they are capable of dealing with (1) imperfectly characterized data and sensor models; and (2) true sensor models and true target models for multisource-multitarget problems. One consequence is that FISST encompasses certain expert-system approaches that are often described as "heuristic"---fuzzy logic, the Dempster-Shafer theory of evidence, and rule-based inference---as special cases of a single probabilistic paradigm.

Section A and Appendix 1 of the report summarize FISST and its basic consequences. Section B summarizes progress made during the course of the contract. Section C summarizes our progress in transitioning this USARO-funded basic research into practical applied-research funded by other DoD agencies.

14. SUBJECT TERMS
data fusion, information fusion, random sets, point processes, multitarget tracking, multitarget statistics, information theory, expert systems

15. NUMBER OF PAGES
66

16. PRICE CODE

17. SECURITY CLASSIFICATION OR REPORT UNCLASSIFIED	18. SECURITY CLASSIFICATION ON THIS PAGE UNCLASSIFIED	19. SECURITY CLASSIFICATION OF ABSTRACT UNCLASSIFIED	20. LIMITATION OF ABSTRACT UL

20030310 021
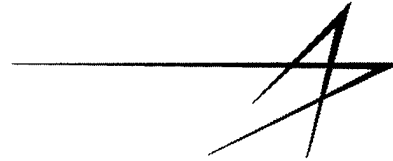
June 28, 2002

"Information-Theoretic Information Fusion"

Final Progress Report

Submitted to: U.S. Army Research Office, Electronics Division

By: Ronald P.S. Mahler, Ph.D.

For: Lockheed Martin Tactical Systems, Eagan MN

In fulfillment of: Contract DAAG55-98-C-0039
(USARO Proposal Number 37629-EL)

Submitted to:
Director
U.S. Army Research Office
ATTN: AMSRL-RO-R1 (DAAD19-00-R-0010)
TPOC: Dr. William Sander
P.O. Box 12211
Research Triangle Park NC 27709-2211

Point of Contact:
Ronald Mahler
Lockheed Martin Tactical Systems
P.O. Box 64525 MS U2S25
St. Paul MN 55164-0525
TEL: 651-456-4819
FAX: 651-456-3098
EML: ronald.p.mahler@lmco.com

FORWARD

INFORMATION-THEORETIC INFORMATION FUSION

This final report for contrct DAAG55-98-C-0039 is submitted to the Electronics Division of the Army Research Office (ARO) by Ronald P.S. Mahler, Ph.D., in behalf of Lockheed Martin Tactical Systems (LMTS) of 3333 Pilot Knob Road, Eagan MN 55121. For the last six years under ARO contracts DAAH04-94-C-0011 and DAAG55-98-C-0039, LMTS has been developing a unified, systematic, and rigorous information-theoretic approach to information fusion. It has been based on "finite-set statistics" (FISST), an "engineering friendly" integration of point process theory and random set theory that was specifically developed under these projects. FISST results in a fully probabilistic unification of detection, classification, tracking, decision-making, sensor allocation, expert-systems theory, situation assessment, and performance evaluation in multi-platform, multi-source, multi-evidence, multi-target, multi-group problems. Highlights are:

(1) A rigorous statistical foundation for multi-sensor, multi-target problems that preserves the practical "Statistics 101" formalism with which engineers are already familiar;

(2) Identifying and correcting several unexpected difficulties in multitarget Bayes-optimal fusion;

(3) Algorithms for simultaneous optimal estimation of numbers, identities, geokinematics of targets;

(4) Systematic foundation for Level 4 fusion (sensor management) based on control theory;

(5) Information-theory based foundation for multi-sensor, multitarget performance evaluation;

(6) Rigorous foundation for detection, tracking, and ID of multiple group targets (force aggregation);

(7) Potentially powerful new computational techniques, based on multitarget statistical analogs of constant-gain Kalman filters (first-order multitarget moment statistics); and

(8) Vigorous technology leverage: The basic research developed under this project is being transitioned into eight applied-research contracts sponsored by agencies such as MRDEC, MDA, AFOSR, three different sites of AFRL, and LMTS.

The primary reporting sections of the report are to be found in sections A, B, and C. In section A we provide an overview of the basic approach, including descriptions of multisensor-multitarget measurement and motion models; the belief-mass functions of a multisensor-multitarget model; the FISST multitarget differential and integral calculus; and unification of expert-systems approaches. Further implications of Finite-Set Statistics can be found in Appendix 1, including: true Bayes-optimal multitarget nonlinear filtering; joint multitarget detection, localization, and identification using multitarget state estimation; unified multi-evidence, multisource, multitarget information fusion; unified multisource-multitarget information theory; multisensor-multitarget sensor management via control theory; and unified multisource-multitarget decision theory. We also address certain published criticisms of FISST in Appendix 2.

Section B summarizes our progress during the contract. This includes: sensor management; measurement models for "ambiguous" evidence; optimal and robust track-track-fusion; computational techniques including a "para-Gaussian" approximation and multitarget first-order moment approximations; algorithmic feasibility analysis; relationship between FISST and point process theory; point target-clusters and continuous-variable finite-set statistics; and multitarget covariance densities and extended Kalman filters.

Section C describes our progress in transitioning our USARO-funded basic research work into applied-research contracts funded by agencies such as MRDEC, AFOSR, AFRL, and MDA.

TABLE OF CONTENTS AND LIST OF APPENDICES

SECTION A: STATEMENT OF PROBLEM STUDIED

A.1 BACKGROUND AND OBJECTIVES

Progress in *single-sensor, single-object* detection, tracking, identification, and information fusion has been greatly facilitated by the existence of a systematic, rigorous, and yet practical engineering statistics that supports the development of new concepts in this field. By "engineering statistics," we mean the vast body of applied mathematical techniques surrounding the following "Statistics 101" concepts that most signal processing engineers learn as undergraduates: (1) random vectors; (2) probability-mass and probability-density functions; (3) differential and integral calculus; (4) statistical moments (expected value, etc.); (5) optimal state estimators; (6) optimal signal-processing filters; and so on. Given the importance of such concepts in the single-sensor, single-object realm, one would expect that multisensor, multitarget information fusion would already rest upon a similarly systematic, rigorous, and yet practical engineering statistics. Surprisingly, until recently this has not been the case. Even more surprisingly, this is true even though a rigorous statistical foundation for multi-object problems—*point process theory* (see section B.8—has been in existence for decades. There appear to be two major reasons for this gap. First and not surprisingly, theoretical development in the multisource-multitarget information fusion community has been focused primarily on immediate engineering applications rather than on systematic, over-arching foundations. Second and perhaps most importantly, neither of the primary mathematical formulations of point process theory—*random measure theory* [13,39,105,121] and *stochastic geometry* (a.k.a. *random set theory*) [4,96,115,131]—have been well-suited for reduction to a practical "Statistics 101" form. What has been missing has been an "engineering friendly" formulation of point process theory—which is to say, one that is *geometric* (in the sense that it treats multi-object systems as visualizable *images*) and which *preserves the "Statistics 101" formalism that signal processing (and especially information fusion) engineers already understand.*

During this and an earlier USARO contract (DAAH04-94-C-0011 and DAAG55-98-C-0039, hereafter described as the "Phase I" and "Phase II" contracts), Lockheed Martin Tactical Systems (LMTS) has developed finite-set statistics (FISST), the "engineering friendly" multi-object statistics that it first introduced into the information fusion community in 1994 [82,84,87,88,89,90]. The core of FISST is a *multisource-multitarget differential and integral calculus* based on the fact that *belief-mass functions* (see section A.2.3) are the multisensor-multitarget counterparts of probability-mass functions. This in turn has led to a solid foundation for multisource-multitarget information theory. The theoretical foundations of FISST have been described in Chapters 2 and 4-8 of the book *Mathematics of Data Fusion* [24], written and published under the Phase I contract. Extended overviews of FISST can be found in the Lockheed Martin technical monograph *An Introduction to Multisource, Multitarget Statistics and Its Applications* [62], the book chapter "Random Set Theory for Target Tracking and Identification" [60], and the short paper "Engineering Statistics for Multi-Object Tracking" [57]. All three were written and published under the Phase II contract.

FISST results in a systematic, fully probabilistic, and statistically rigorous unification of detection, classification, tracking, decision-making, sensor allocation, situation assessment, expert-systems theory, and performance evaluation in multi-platform, multi-source, multi-evidence, multi-target, multi-group problems. Highlights are:

(1) A rigorous basis for multisource-multitarget information theory;
(2) A unified, rigorous, and probabilistic foundation for many aspects of expert systems theory (fuzzy logic, the Dempster-Shafer theory, rule-based evidence, Bayesian statistics);
(3) Identifying and addressing several unexpected difficulties in multitarget Bayes-optimal tracking;

4

(4) Algorithms for simultaneous optimal estimation of numbers, identities, geokinematics of targets;

(5) Systematic foundation for sensor management based on control theory;

(6) Information theory-based foundation for multisensor, multitarget performance evaluation;

(7) Rigorous foundation for detection, tracking, and ID of multiple group targets (force aggregation);

(8) Potentially powerful new computational techniques, based on multi-target statistical analogs of constant-gain Kalman filters (first-order multitarget moment statistics);

(9) Vigorous technology leverage: The basic research developed under this project is being transitioned into eight applied-research contracts sponsored by agencies such as MRDEC, MDA, AFOSR, and three different sites of AFRL, and LMTS; and

(10)   Widespread interest in FISST techniques. For example, project PI Dr. Ronald Mahler has gave an invited to give a two-day tutorial at the International Conference on Information, Decision, and Control at the University of Adelaide, Australia, in mid-February 2002. A group at the University of Melbourne has been given a grant to study random set methods. A member of this team, Dr. Ba-Ngo Vo, spent 1½ weeks with Dr. Mahler in August 2001 to learn more about FISST techniques. Shorter versions of this tutorial are to be given at the 2002 International Conference on Information Fusion; and the 2002 IEEE Workshop on Multi-Object Tracking.

Overall, the response to FISST in the information fusion community has been very positive. In particular, this was the case with the peer reviews for our second contract. (There have been a few published criticisms of FISST, which will be addressed in Appendix 2 below.) However, these reviewers also strongly recommended that it was time to put FISST to the test by applying it to real-world problems. While justified, this criticism presented LMTS with a quandary. The budget under a basic research contract stretches only so far, and concentrating on a single "pet rock technology" runs the risk of squandering scarce resources on a solution that nobody actually wants, despite all expectations to the contrary. Consequently, LMTS addressed the reviewers' recommendation in an unusual manner. Leveraging our Phase I and Phase II USARO contracts as basic-research "intellectual venture capital," we used FISST to develop innovative techniques directed at a wide range of information fusion applications. Our belief was that at least some of these would attract enough funding to support application to real-world problems. This "omnidirectional" technology-leveraging strategy has proved very successful. While some of the techniques developed under the first two contracts have not attracted significant funding attention as yet, several others have. As a result, FISST-based techniques are currently being investigated in a range of real applications (many using real data) funded by a number of DoD agencies, including AFOSR, AFRL/IFEA, AFRL/SNAT, MDA, and MRDEC. These include:

(1) scientific multisource-multitarget performance estimation based on information theory;

(2) cluster target tracking and discrimination for ballistic missile defense;

(3) robust target identification fusion using multisource High Range Resolution Radar (HRRR);

(4) robust automatic target recognition against ground targets using Synthetic Aperture Radar (SAR);

(5) fundamental fusion and control technologies for swarms of UCAVs.

See section C below for more details.

In the remainder of this section, we summarize the basic elements of FISST and the progress made during the second of our two previous USARO contracts. In section A.2 we provide an overview of FISST, including the following topics:

(1) description of the basic approach (A.2.1);

(2) multisensor-multitarget measurement and motion models (A.2.2);

(3) belief-mass functions of a multisensor-multitarget measurement or motion model (A.2.3);

(4) the FISST multitarget differential and integral calculus (A.2.4); and

(5) the FISST unification of expert-systems approaches, e.g. fuzzy logic, Dempster-Shafer, etc. (A.2.5)

5

In Appendix 1 we summarize some important consequences of FISST:

(1) true Bayes-optimal multitarget nonlinear filtering (1-1);
(2) joint multitarget detection and estimation (1-2);
(3) unified multi-evidence, multisource, multitarget information fusion (1-3);
(4) unified multisource-multitarget information theory, including Cramér-Rao bounds (1-4);
(5) sensor management via multisource-multitarget control theory (1-5); and
(6) unified multisource-multitarget decision theory (1-6);

In Appendix 2, we describe and address some recent published criticisms of FISST.

In section B, we describe progress made during the Phase II contract:

(1) progress in sensor management analysis (B.1);
(2) progress in scientific performance estimation (B.2);
(3) progress in measurement models for ambiguous evidence (B.3);
(4) progress in Level 2 information fusion, a.k.a. Situation Assessment (B.4);
(5) progress in track-to-track fusion (B.5);
(6) progress in computational techniques, including first-order multitarget moment filters and a multitarget "para-Gaussian" approximation (B.6);
(7) progress in algorithmic feasibility analysis (B.7);
(8) unplanned progress: relationship between FISST and conventional point process theory (B.8);
(9) unplanned progress: continuous-state multitarget statistics (B.9); and
(10) unplanned progress: examination of concept of "multitarget extended Kalman filter" (B.10).

In section B.11 we describe progress in transitioning technology developed under the Phase II contract into practical application.

A.2 AN OVERVIEW OF FINITE-SET STATISTICS

A.2.1 The Basic Approach. The basic approach is as follows. Suppose that a suite of known sensors (or other information sources) interrogates multiple targets (whose number, positions, velocities, identities, threat states, etc. are all unknown) and transmits all observations to a central data fusion site. Then FISST is based on the following sequence of ideas [24,60,62,74,87,88,90]:

(1) reconceptualize all sensors as a single sensor;

(2) reconceptualize the randomly varying set of targets, of randomly varying number n, as a single target with multitarget state-set $X = \{x_1,..., x_n\}$;

(3) reconceptualize the set $Z = \{z_1,..., z_m\}$ of random observations of randomly varying number m, collected by the sensor suite at approximately the same time, as a single measurement of the target-set observed by the sensor suite;

(4) just as single-sensor, single-target data (without missed detections or false alarms) can be modeled using a measurement model $Z_k = h_k(x, W_k)$, so model multisensor, multitarget data using a *multisensor-multitarget measurement model*—i.e., a randomly varying finite set $\Sigma_k = T_k(X) \cup C_k(X)$, where $T_k(X)$ indicates target-generated observations and $C_k(X)$ indicates clutter-generated observations (section A.2.2);

(5) just as single-target motion can be modeled using a motion model $\mathbf{X}_{k+1} = \Phi_k(\mathbf{x}_k, \mathbf{V}_k)$, model the motion of multitarget systems using a *multitarget motion model*—i.e., a randomly varying finite set $\Gamma_{k+1} = \Phi_k(X_k) \cup B_k(X_k)$, where $\Phi_k(X_k)$ indicates the time-evolution of old targets (some of which may disappear) and where $B_k(X_k)$ indicates the appearance of new "birth" targets (section A.2.2);

(6) given "ambiguous data" (e.g. natural language reports, datalink and other features produced by human interpretation, knowledge-base rules, etc.), model such data as random subsets Θ of measurement space, and the statistics of such data as "generalized likelihood functions" $\rho(\Theta | \mathbf{x})$ (see section B.3.1);

Given this, we can *reformulate multisensor, multitarget problems as abstract single-sensor, single-target problems*. The basis of this reformulation is *belief-mass*. Belief-mass functions are non-additive generalizations of probability-mass functions. (Nevertheless, they are not heuristic: they are equivalent to probability-mass functions on certain abstract topological spaces. See section A.2.3) That is:

(7) just as the probability-mass function $p_k(S|\mathbf{x}) = \Pr(\mathbf{Z}_k \in S)$ of a single-sensor, single-target measurement model model $\mathbf{Z}_k = \mathbf{h}_k(\mathbf{x}, \mathbf{W}_k)$ describes the statistics of single-sensor data, the belief-mass function $\beta_k(S|X) = \Pr(\Sigma_k \subseteq S)$ of a multisource-multitarget measurement model set $\Sigma_k = T_k(X) \cup C_k(X)$ describes the statistics of multisource-multitarget data (section A.2.3);

(8) just as the probability-mass function $p_{k+1|k}(S|\mathbf{x}) = \Pr(\mathbf{X}_{k+1} \in S)$ of a single-target motion model model $\mathbf{X}_{k+1} = \Phi_k(\mathbf{x}_k, \mathbf{V}_k)$ is used to describe the statistics of single-target motion, use the belief-mass function $\beta_{k+1|k}(S|X) = \Pr(\Gamma_{k+1} \subseteq S)$ of a multitarget motion model $\Gamma_{k+1} = \Phi_k(X_k) \cup B_k(X_k)$ to describe the statistics of multitarget motion (section A.2.3).

The FISST *multisensor-multitarget differential and integral* is what transforms these mathematical abstractions into practical form:

(9) Just as the true single-sensor, single-target likelihood function $f_k(\mathbf{z}|\mathbf{x})$ can be derived from the belief-mass function $p_k(S|\mathbf{x})$ of the single-sensor, single-target measurement model via differentiation, so the true multisensor-multitarget likelihood function $f_k(Z|X)$ can be derived from the belief-mass function $\beta_k(S|X)$ of the multisensor-multitarget measurement model using a generalized differentiation operator called the *set derivative* (section A.2.4);

(10)   Just as the true Markov transition density $f_{k+1|k}(\mathbf{y}|\mathbf{x})$ can be derived from the probability-mass function $p_{k+1|k}(S|\mathbf{x})$ of the single-target motion model via differentiation, so the true multitarget Markov transition density $f_{k+1|k}(Y|X)$ can be derived from the belief-mass function $\beta_{k+1|k}(S|X)$ of the multitarget motion model via set-differentiation (section A.2.4);

(11)   Just as the density $f_k(\mathbf{z}|\mathbf{x})$ and the probability-mass function $p_k(S|\mathbf{x})$ are related by the equation $p_k(S|\mathbf{x}) = \int_S f_k(\mathbf{z}|\mathbf{x})d\mathbf{z}$, so the multi-object density $f_k(Z|X)$ and the belief-mass function $\beta_k(S|X)$ are related by the equation $\beta_k(S|X) = \int_S f_k(Z|X)\delta Z$, where the integral is now a multisource-multitarget *set integral* (section A.2.4).

Given this, let $Z^{(k)} = \{Z_1, ..., Z_k\}$ be a time-sequence of multisource-multitarget observations. Then one can construct *true multitarget posterior distributions* from the true multisource-multitarget likelihood, using Bayes' rule: $f_{k+1|k+1}(X| Z^{(k+1)}) \propto f_{k+1}(Z_{k+1}|X) f_{k+1|k}(X| Z^{(k)})$ [90, p.337]. Here,

$$f_{k|k}(\varnothing| Z^{(k)}) = \text{posterior likelihood of no targets}$$

$$f_{k|k}(\{\mathbf{x}\}|Z^{(k)}) = \text{posterior likelihood of one target with state } \mathbf{x}$$

$$f_{k|k}(\{\mathbf{x}_1,\mathbf{x}_2\}|Z^{(k)}) = \text{posterior likelihood of } n \text{ targets with states } \mathbf{x}_1,...,\mathbf{x}_n$$

$$\vdots$$

$$f_{k|k}(\{\mathbf{x}_1,...,\mathbf{x}_n\}|Z^{(k)}) = \text{posterior likelihood of } n \text{ targets with states } \mathbf{x}_1,...,\mathbf{x}_n$$

From these distributions one can in turn compute simultaneous, provably optimal estimates of target number, kinematics, and identity without resort to the optimal report-to-track assignment characteristic of multi-hypothesis approaches. We also have the means of accomplishing both *optimal-Bayes* and *robust-Bayes* multisensor-multitarget information fusion, detection, tracking, and identification.

Random Vector, Z	Random Finite Set, Σ
sensor,	"meta-sensor,"
target,	"meta-target,"
observation, $\mathbf{z}$	observation-set, Z
state vector, $\mathbf{x}$	state-set, X
measurement model, $\mathbf{Z}_k = \mathbf{h}_k(\mathbf{x},\mathbf{W}_k)$	multitarget m.m, $\Sigma_k = T_k(X) \cup C_k(X)$
motion model, $\mathbf{X}_{k+1} = \Phi_k(\mathbf{x}_k,\mathbf{V}_k)$	multitarget m.m., $\Gamma_{k+1} = \Phi_k(X_k) \cup B_k(X_k)$
differentiation, $dp_k/d\mathbf{z}$	set differentiation, $\delta\beta_k/\delta Z$
integration, $\int f_k(\mathbf{z}|\mathbf{x})d\mathbf{z}$,	set integration, $\int f_k(Z|X)\delta Z$
probability-mass function, $p_k(S|\mathbf{x})$	belief-mass function, $\beta_k(S|X)$
likelihood function, $f_k(\mathbf{z}|\mathbf{x})$	multitarget likelihood function, $f_k(Z|X)$
Markov density, $f_{k+1|k}(\mathbf{y}|\mathbf{x})$	multitarget Markov density, $f_{k+1|k}(Y|X)$
posterior density, $f_{k|k}(\mathbf{x}|Z^k)$.	multitarget posterior density, $f_{k|k}(X|Z^{(k)})$.
recursive Bayes filtering	recursive multitarget Bayes filtering
information theory	multisource-multitarget information theory
miss distance	multitarget miss distance
control theory	multisensor-multitarget sensor management

It thus turns out that we get a list of *direct mathematical parallels* between the world of single-sensor, single-target statistics and the world of multisensor, multitarget statistics, as illustrated in the above table. This parallelism is so close that general statistical methodologies can, with a bit of prudence, be directly "translated" from the single-sensor, single-target case to the multisensor-multitarget case. That is, the table can be thought of as a "dictionary" that establishes a direct correspondence between the words and grammar in the random-vector language and cognate words and grammar of the random-set language. Consequently, any "sentence" (any concept or algorithm) phrased in the random-vector language can, in principle, be directly "translated" into a corresponding sentence (corresponding concept or algorithm) in the random-set language. This process can be encapsulated as a general methodology for attacking multisource-multitarget data fusion problems that has been the fundamental motivating philosophy behind FISST since 1994:

> **Almost-Parallel Worlds Principle (APWOP):** *Nearly any concept or algorithm phrased in random-vector language can, in principle, be directly translated into a corresponding concept or algorithm in the random-set language.* **[87,90,62]**

We say "almost-parallel" because, as with any translation process, the correspondence between dictionaries is not precisely one-to-one (for example, vectors can be added and subtracted whereas finite sets cannot). Nevertheless, the parallelism is complete enough that, provided one exercises some care, a hundred years of accumulated knowledge about single-sensor, single-target statistics can be *directly* brought to bear on multisensor-multitarget problems.

The following simple example has been used to illustrate the APWOP since 1994. The performance of a multitarget data fusion algorithm can be measured by constructing *information-based measures of effectiveness*, e.g. the following multitarget generalization of the Kullback-Leibler discrimination **[24,87,90,83]**

Single-sensor, single-target	$\Rightarrow$	Multisensor-multitarget
$K(f;g) = \int f(\mathbf{x}) \log\left(\dfrac{f(\mathbf{x})}{g(\mathbf{x})}\right) d\mathbf{x}$	$\Rightarrow$	$K(f;g) = \int f(X) \log\left(\dfrac{f(X)}{g(X)}\right) \delta X$

Here we have used the APWOP to replace conventional statistical concepts with their FISST multisensor, multitarget counterparts. The ordinary densities *f, g* on the left are replaced by the multitarget densities *f, g* on the right; and the ordinary integral on the left is replaced by a FISST set integral on the right. References **[24, pp. 295-312]** and **[33,69,140,141]** of section G below describe how this application of the APWOP leads to a systematic approach to to scientific performance estimation for multisensor, multitarget algorithms.

Items (1)-(10) above are summarized in somewhat greater detail in the following subsections.

A.2.2 Multisensor-Multitarget Measurement and Motion Models. It is possible to construct measurement models for multisensor-multitarget problems in much the same way that one constructs measurement models for single-sensor, single-target problems, as indicated in the following table:

measurement model:	$\mathbf{Z}_k = \mathbf{h}_k(\mathbf{x}, \mathbf{W}_k)$
multisensor-multitarget measurement model:	$\Sigma_k = T_k(X) \cup C_k(X)$

where $X = \{\mathbf{x}_1, ..., \mathbf{x}_n\}$ is the random multitarget state-set and where $C_k(X)$ models false alarms and/or (possibly state-dependent) point clutter. As an example, drop the index k and assume that $C(X)$ has the form $C = C_1 \cup ... \cup C_m$ where each C_j is a state-independent clutter generator—meaning that there is a probability p_{FA} that C_j will be non-empty (i.e., generate a clutter-observation). In this case $C_j = \{\mathbf{C}_j\}$ where $\mathbf{C}_j$ is a random noise vector with density $f_C(\mathbf{z})$.

In like manner, it is possible to construct multitarget motion models in much the same way that one constructs motion models for single-target problems:

motion model:	$\mathbf{X}_{k+1} = \Phi_k(\mathbf{x}, \mathbf{V}_k)$
multitarget motion model:	$\Gamma_{k+1} = \Phi_k(X) \cup B_k(X_k)$

For example, let $X = \varnothing$, $X = \{\mathbf{x}\}$, or $X = \{\mathbf{x}_1, \mathbf{x}_2\}$ (i.e., no more than targets in the scene). Also let

$$\Gamma_{k+1} = \varnothing$$

$$\Gamma_{k+1} = T_{k,\mathbf{x}}$$

$$\Gamma_{k+1} = T_{k,\mathbf{x}_1} \cup T_{k,\mathbf{x}_2}$$

where $T_{k,\mathbf{x}}$ is a track-set with the following properties: (a) $T_{k,\mathbf{x}} \neq \varnothing$ with probability p_V, in which case $T_{k,\mathbf{x}} = \{\mathbf{X}_{k,\mathbf{x}}\}$; and (b) $T_{k,\mathbf{x}} = \varnothing$ (i.e., target disappearance), with probability $1 - p_V$. In other words, if no targets are present in the scene then this will continue to be the case. If there is one target in the scene then either this target will persist (with probability p_V) or it will vanish (with probability $1 - p_V$). If there are two targets in the scene, then each will either persist or vanish.

A.2.3 Belief-Mass Functions of a Multisensor-Multitarget Model. Just as the statistical behavior of a random observation vector $\mathbf{Z}_k$ is characterized by its probability mass function $p_k(S|\mathbf{x}) = \Pr(\mathbf{Z}_k \in S)$, so the statistical behavior of the random observation-set Σ_k is characterized by its *belief-mass function*

$$\beta_k(S|X) = \Pr(\Sigma_k \subseteq S)$$

The belief mass is just the *total probability that all observations in a sensor (or multi-sensor) scan* will be found in any given region S, if targets have multitarget state X. For example, if $X = \{\mathbf{x}\}$ and $\Sigma_k = \{\mathbf{Z}_k\}$ where $\mathbf{Z}_k$ is a random vector then

$$\beta_k(S|X) = \Pr(\Sigma_k \subseteq S) = \Pr(\mathbf{Z}_k \in S) = p_k(S|\mathbf{x})$$

In other words, the belief mass of a random vector is just its probability mass.

Likewise, in single-target problems the statistics of a motion model $\mathbf{X}_{k+1} = \Phi_k(\mathbf{x}_k, \mathbf{V}_k)$ are described by the probability-mass function $p_{k+1|k}(S|\mathbf{x}) = \Pr(\mathbf{X}_{k+1} \in S)$, which is the probability that the target-state will be found in the region S if it previously had state $\mathbf{x}_k$. Suppose that $\Gamma_{k+1} = \Phi_k(X_k) \cup B_k(X_k)$ is a multitarget motion model. Then the statistics of the finitely varying random state-set Γ_{k+1} is described by its belief-mass function

$$\beta_{k+1|k}(S|X) = \Pr(\Gamma_{k+1} \subseteq S)$$

This is the total probability of finding all targets in region S at time-step $k+1$ if, in time-step k, they had multitarget state $X_k = \{\mathbf{x}_{k,1}, \dots, \mathbf{x}_{k,n(k)}\}$.

Note: The concept of a belief-mass function is not *ad hoc* but, rather, derives directly from the standard concept of a probability-mass function (a.k.a. probability measure). One begins with an abstract space whose points are finite sets of objects drawn from some other space (typically, a single-target state space or single-sensor, single-target measurement spaces). There are many possible ways to topologize such spaces [97]. FISST uses the Mathéron "hit-or-miss" topology [96] for three reasons. First, the Mathéron topology can be thought of as the simplest way of extending conventional Euclidean topology to spaces whose points are the closed subsets of Euclidean measurement or state spaces [24, pp. 131-135]. Second, it allows us to subsume Bayesian probability, the Dempster-Shafer theory, and fuzzy logic under a common probabilistic paradigm. Second, this topology can be safely ignored for purposes of application, since it transforms *probability masses on abstract subset-spaces* into *belief-masses on ordinary Euclidean spaces*. Specifically, let Σ be a random subset, O any Borel subset of the Mathéron topology, and $p_\Sigma(O) = \Pr(\Sigma \in O)$ the probability-mass function of Σ concentrated on O. Then the Choquet-Mathéron capacity theorem [96, p. 30] tells us that the additive probability function p_Σ is uniquely determined by the specific *non-additive* set function

$$\beta_\Sigma(S) = \Pr(\Sigma \subseteq S) = \Pr(\Sigma \in O_{S^c}^c) = p_\Sigma(O_{S^c}^c)$$

where $O = O^c_{S^c}$ denotes the class whose elements are all closed subsets C (of the underlying measurement or state space) such that $C \cap S^c = \emptyset$ (i.e., $C \subseteq S$) where S is some closed subset of the underlying space.

A.2.4 The FISST Calculus. Let $f(X)$ be a nonnegative-valued function of a variable X that ranges over finite subsets of objects in some space of interest (e.g. multitarget states, multisensor-multitarget observations, etc.). Then the *set integral* of $f(X)$ is

$$\int_S f(X)\delta Z = f_S(0) + f_S(1) + f_S(2) + \dots + f_S(n) + \dots$$

where $f_S(0)$ is the probability of there being no objects in S, and where

$$f_S(n) = \frac{1}{n!} \int_{S \times .. \times S} f(\mathbf{x}_1,...,\mathbf{x}_n) d\mathbf{x}_1 \cdots d\mathbf{x}_n$$

is the marginal probability that there are $n = 0,1,2,...$ objects present in S. (Set integrals appear regularly in the statistical theory of gases and liquids, though they are not explicitly identified as such in that context [**31, pp. 234, eq. 37.4; 266, eq. 40.28**].) Write $\int f(X)\delta Z = \int_S f(X)\delta Z$ if S is the entire space.

The inverse operation of the set integral, the *set derivative*, is a generalization of the so-called Radon-Nikodým derivative of real analysis. That is, let $\beta(S)$ be any function whose arguments S are arbitrary closed subsets. (Typically, β will be the belief-mass function of a multisensor-multitarget measurement model or of a multitarget motion model.) If $Z = \{\mathbf{z}_1,..., \mathbf{z}_m\}$ with $\mathbf{z}_1,..., \mathbf{z}_m$ distinct, define:

$$\frac{\delta\beta}{\delta\mathbf{z}}(S) = \lim_{\lambda(E_\mathbf{z}) \to 0} \frac{\beta(S \cup E_\mathbf{z}) - \beta(S)}{\lambda(E_\mathbf{z})}$$

$$\frac{\delta\beta}{\delta Z}(S) = \frac{\delta}{\delta\mathbf{z}_1} \cdots \frac{\delta}{\delta\mathbf{z}_m} \beta(S)$$

$$\frac{\delta\beta}{\delta\emptyset}(S) = \beta(S)$$

(*Caution*: The first of these three equations is simplified for clarity. A more complicated definition is required to encompass discrete variables and situations in which $S \cap E_\mathbf{z} \neq \emptyset$ and to ensure that the limit is not ill-defined.) . The set derivative can also be thought of as a special kind of functional derivative (see section B.8).

The importance of the set derivative arises, in part, from the fact that it can be used to explicitly construct multisensor-multitarget likelihood functions and multitarget Markov densities. That is,

- The *true likelihood function* $f_k(Z|X)$ of a multisensor-multitarget problem is a set derivative of the belief mass function $\beta_k(S|X)$ of the corresponding sensor (or multi-sensor) model:

$$f_k(Z \mid X) = \frac{\delta\beta_k}{\delta Z}(\emptyset \mid X)$$

That is, it is "true" in the sense that $\int_S f_k(Z|X)\delta Z = \beta_k(S|X)$, where the integral is a set integral.

- The *true Markov transition density* $f_{k+1|k}(Y|X)$ of a multitarget problem is a set derivative of the belief mass function $\beta_{k+1|k}(S|X)$ of the corresponding multitarget motion model:

$$f_{k+1|k}(Y \mid X) = \frac{\delta \beta_{k+1|k}}{\delta Y}(\varnothing \mid X)$$

Once again, it is "true" since $\int_S f_{k+1|k}(Y|X)\delta Y = \beta_{k+1|k}(S|X)$, where the integral is a set integral.

Let S_{tot} be the complete (single-target) state space. Then in addition, the set derivative can be used to compute *multitarget statistical moments* (see sections B.6.2, B.8, and B.10.1) of a multitarget posterior:

• The *multitarget moment density* of a multisensor-multitarget problem is a set derivative of the belief-mass function $\beta_{k|k}(S|X)$ of the corresponding multitarget posterior:

$$m_{k|k}(X \mid Z^{(k)}) = \frac{\delta \beta_{k|k}}{\delta X}(S_{tot} \mid X)$$

• The *multitarget covariance density* of a multisensor-multitarget problem is a set derivative of the logarithm of the belief-mass function $\beta_{k|k}(S|X)$ of the corresponding multitarget posterior:

$$c_{k|k}(X \mid Z^{(k)}) = \frac{\delta \log \beta_{k|k}}{\delta X}(S_{tot} \mid X)$$

Because set derivatives are defined in terms of complicated limits, they might seem to merely transform one difficult problem—constructing multitarget likelihood functions and Markov densities—into another equally difficult problem—computing very complex limits. However, ``turn the crank" rules exist for the FISST calculus, e.g. the following sum, product, chain, and power rules **[24, p.151]**, **[62, pp. 31-32]**:

Sum Rule:

$$\frac{\delta}{\delta Z}[a_1\beta_1(S) + a_2\beta_2(S)] = a_1\frac{\delta \beta_1}{\delta Z}(S) + a_2\frac{\delta \beta_2}{\delta Z}(S)$$

Product Rules:

$$\frac{\delta}{\delta z}[\beta_1(S)\beta_2(S)] = \beta_1(S)\frac{\delta \beta_2}{\delta z} + \frac{\delta \beta_1}{\delta z}\beta_2(S)$$

$$\frac{\delta}{\delta Z}[\beta_1(S)\beta_2(S)] = \sum_{W \subseteq Z}\frac{\delta \beta_1}{\delta W}(S)\frac{\delta \beta_2}{\delta(Z-W)}(S)$$

Chain Rule:

$$\frac{\delta}{\delta z}f(\beta_1(S),...,\beta_n(S)) = \sum_{i=1}^{n}\frac{\partial f}{\partial x_i}(\beta_1(S),...,\beta_n(S))\frac{\delta \beta_i}{\delta z}(S)$$

Power Rule:

$$\frac{\delta}{\delta Z}p(S)^n = \begin{cases} \dfrac{n!}{(n-k)!}p(S)^{n-k}f_p(z_1)\cdots f_p(z_k) & (k \leq n) \\[4mm] 0 & (k > n) \end{cases}$$

where in the last equation, $Z = \{z_1,...,z_k\}$, $n \geq 0$ is an integer, and $p(S)$ is a probability mass function with density function $p_f(x)$.

A.2.5 Unification of Expert-Systems Approaches. One of the more novel features of FISST is the fact that it subsumes, under a single probabilistic paradigm (and therefore one that is compatible with Bayesian techniques) many expert-systems approaches, e.g. fuzzy logic, the Dempster-Shafer theory of evidence, and rule-based inference. It is based on the notion that *ambiguous data can be probabilistically represented as random closed subsets of (multisource) measurement space.*

Consider a simple example (for a more extensive discussion see [24, pp. 266-269]). Suppose that we are given the sensor measurement model $Z = Cx + W$ where x is the target state, W is random noise, and C is an invertible matrix. Let B be an observation that is "imprecise" in the sense that it is a subset of measurement space that merely constrains the possible values of z—i.e., $z \in B$. Then the random variable Γ defined by $\Gamma = \{C^{-1}(z\text{-}W) \mid z \in B\}$ is the randomly varying subset of all target states that are consistent with this imprecise observation. That is, the imprecise observation B *indirectly* constrains the possible target *states* as well.

Suppose, more generally, that we are not very certain about the validity of the constraint $z \in B$ but, rather, believe that there are many possible constraints—of varying plausibility—on z. Then we would model this kind of ambiguity as a randomly varying subset Θ of measurements, where the probability $Pr(\Theta = B)$ represents our degree of belief in the specific constraint B. The random subset of all states that are consistent with Θ would then be $\Gamma = \{C^{-1}(z\text{-}W) \mid z \in \Theta\}$. (*Caution*: The random closed subset Θ is a model of *a single observation collected by a single source.* It should not be confused with a *multisensor, multitarget* observation-set Σ, whose instantiations $\Sigma = Z$ are finite random sets that have the general form $Z = \{z_1, ..., z_m, \Theta_1, ..., \Theta_{m'}\}$ where $z_1, ..., z_m$ are individual conventional observations and $\Theta_1, ..., \Theta_{m'}$ are random-set models of individual ambiguous observations.)

It is one thing to recognize that random sets provide a common probabilistic foundation for various kinds of statistically ill-characterized data. It is quite another to construct practical random set representations of such data. The following paragraphs show how three kinds of ambiguous data— imprecise, vague, and contingent—can be represented probabilistically by random sets [27,49,75,80,113] and how these, in turn, can be modeled using likelihood functions.

A.2.5-1 *Vague data: fuzzy logic.* A fuzzy membership function on some (finite or infinite) universe U is a function that assigns a number $f(u)$ between zero and one to each member u of U. The random subset $\Theta = \Sigma_A(f)$, called the *canonical random set representation* of the fuzzy subset f, is defined by

$$\Sigma_A(f) = \{u \in U \mid A \leq f(u)\}$$

where A is a uniformly generated random number on the unit interval [0,1]. [21,22,25,34,108,109]

A.2.5-2 *Imprecise data: Dempster-Shafer bodies of evidence.* A Dempster-Shafer body of evidence B on some universe U consists of a list $B_1, ..., B_b$ of nonempty subsets of U and nonnegative weights $b_1, ..., b_b$ that sum to one. Let Θ be a random subset of U such that $Pr(\Theta = B_j) = b_j$ for $j = 1, ..., b$. Then Θ is the random set representation of B and we write $B = B^\Theta$ [30,107,118]. The Dempster-Shafer theory can be generalized to the case when the B_j are *fuzzy* subsets of U [81]. Such "fuzzy bodies of evidence" can also be represented in random set form. A Bayes-compatible version of the Dempster-Shafer rule of combination can also be defined [18,76,81,86].

A.2.5-3 *Contingent data: rule-based inference and conditional event algebra.* Knowledge-base rules have the form $X \Rightarrow S = $ "*if X then S*" where X,S subsets of a (finite) universe U with N elements. Using the Goodman-Nguyen theory of conditional event algebras [23], LMTS has shown [78,79] that there is at least one way to represent knowledge-base rules in random set form. Specifically, let Φ a

uniformly distributed random subset of U—that is, one whose probability distribution is $\Pr(\Phi = S) = 2^{-N}$ for all subsets S of U. A random set representation $\Theta = \Sigma_\Phi(X\Rightarrow S)$ of the rule $X \Rightarrow S$ is:

$$\Sigma_\Phi(X \Rightarrow S) = (S \cap X) \cup (X^c \cap \Phi)$$

See section B.3.1 below for a description of how one can construct *generalized likelihood functions* that describe the informativeness of such data.

A.2.5-4 *General random set models.* A mathematical construction due to Y. Li [54] provides a means of easily constructing quite general random set models of evidence of any universe U. See **[24, pp. 265-266]** or **[62, p. 60]** for details.

A.3 CONSEQUENCES OF FINITE-SET STATISICS

The previous sections summarized what FISST "is." The following subsections of Appendix 1 below summarize some fundamental consequences of FISST:

(1) true Bayes-optimal multitarget nonlinear filtering (1-1);
(2) joint multitarget detection, localization, and identification (section 1-2);
(3) unified multi-evidence, multi-source, multi-target information fusion (1-3);
(4) unified multisource-multitarget information theory, with multitarget Cramér-Rao bounds (1-4);
(5) sensor management based on unified multisource-multitarget control theory (1-5); and
(6) unified multisource-multitarget decision theory and ROC curves (1-6).

14

SECTION B: SUMMARY OF MOST IMPORTANT RESULTS

In what follows, we summarize the technical progress completed under the Phase II contract. (Progress in terms of technology transition and academic publication is reported in sections C and D, respectively.) Specifically, we describe:

(1) progress in sensor management analysis (B.1);
(2) progress in scientific performance estimation (B.2);
(3) progress in measurement models for ambiguous evidence (B.3);
(4) progress in Level 2 information fusion, a.k.a. Situation Assessment (B.4);
(5) progress in track-to-track fusion (B.5);
(6) progress in computational techniques, including first-order multitarget moment filters, multitarget particle-system filters, and a multitarget "para-Gaussian" approximation (B.6);
(7) progress in algorithmic feasibility analysis (B.7);
(8) unplanned progress: relationship between FISST and point process theory (B.8);
(9) unplanned progress: continuous-state multitarget statistics (B.9); and
(10) unplanned progress: examination of concept of "multitarget extended Kalman filter" (B.10).

B.1 PROGRESS IN SENSOR MANAGEMENT ANALYSIS

In the proposal for the Phase II contract, we suggested further investigation of the control-theoretic approach sketched out during the Phase I contract. As noted in section 1-5 of Appendix 1, the basic goal of a multisensor-multitarget sensor management system is to choose the latest control set $\mathbf{u}_k$ to maximize the "peakiness" of the following likelihood ratio

$$r_{Z_{k+1}}(X, \vec{\mathbf{x}}^* \mid \mathbf{u}_k) = \frac{f_k(Z_{k+1} \mid X, \vec{\mathbf{x}}^*)}{f_k(Z_{k+1} \mid Z^{(k)}, U^{k-1}, \mathbf{u}_k)}$$

independently of what actual multisensor-multitarget observation-set Z_{k+1} might be collected next. During the Phase II contract, LMTS made some progress towards reaching a better understanding of this problem. First one must construct a measure of "peakiness" of the distribution. This can be accomplished in a number of ways. For example, we can try to maximize the supremal likelihood ratio

$$r_{Z_{k+1}}(\mathbf{u}_k) = \sup_{X, \vec{\mathbf{x}}^*} r_{Z_{k+1}}(X, \vec{\mathbf{x}}^* \mid \mathbf{u}_k)$$

or we can try to maximize the average log-likelihood ratio:

$$r_{Z_{k+1}}(\mathbf{u}_k) = E\left[\log r_{Z_{k+1}}(\Xi, \vec{\mathbf{X}}^* \mid \mathbf{u}_k)\right] = \int \log r_{Z_{k+1}}(X, \vec{\mathbf{x}}^* \mid \mathbf{u}_k) f_{k+1|k+1}(X \mid Z^{(k)}, Z_{k+1}, \mathbf{u}_k) \delta X d\vec{\mathbf{x}}^* \text{ or the}$$

Then, one must hedge against the fact that the next observation-set Z_{k+1} cannot be known ahead of time. This can also be accomplished in a number of ways. We can hedge against the worst case if we use a minimax approach

$$r(\mathbf{u}_k) = \inf_{Z_{k+1}} r_{Z_{k+1}}(\mathbf{u}_k)$$

or we can instead hedge against the average case

$$r(\mathbf{u}_k) = E\left[r_{\Xi_{k+1}}(\mathbf{u}_k)\right] = \int r_{Z_{k+1}}(\mathbf{u}_k) f(Z_{k+1} \mid Z^{(k)}) \delta Z_{k+1}$$

In our Phase I contract, we hedged using the "non-informative observation" $Z_{k+1} = \varnothing$:

$$r(\mathbf{u}_k) = r_\varnothing(\mathbf{u}_k)$$

This approach can be regarded as an approximation to the minimax procedure. In minimax, we determine the observation $Z_{k+1} = Z_*$ that produces worst-case peakiness. Even if $Z_* \neq \emptyset$, the observation $Z_{k+1} = \emptyset$ is very much like a worst-case observation: It will most typically occur because $\mathbf{u}_k$ has been chosen so poorly that no target is in the Field of View of any sensor.

It is unclear at this time which of the many possible objective functions is the best to use, whether from a computational or a performance point of view. As an example, however, measure peakiness by maximizing over state variables and hedge against unknown observations using the non-informative observation. In this case we get

$$r_\emptyset(\mathbf{u}_k) = \sup_{X,\vec{\mathbf{x}}^*} r_\emptyset(X,\vec{\mathbf{x}}^* \mid \mathbf{u}_k) = \frac{\sup_{X,\vec{\mathbf{x}}^*} f_k(\emptyset \mid X,\vec{\mathbf{x}}^*)}{f_k(\emptyset \mid Z^{(k)},U^{k-1},\mathbf{u}_k)}$$

and so maximizing $r_\emptyset(\mathbf{u}_k)$ is the same thing as minimizing

$$f_k(\emptyset \mid Z^{(k)},U^{k-1},\mathbf{u}_k) = \int f_{k+1}(\emptyset \mid X) f_{k+1|k}(X \mid Z^{(k)})\delta X$$

That is, we minimize the average value of the probability $f_{k+1}(\emptyset|X)$ of *not detecting any target at all*. Consequently, the probability $1 - f_{k+1}(\emptyset|X)$ (of collecting at least one observation) behaves somewhat like a "multitarget probability of detection." This work has been greatly extended under our new AFOSR contract (section C.4).

B.2 PROGRESS IN SCIENTIFIC PERFORMANCE ESTIMATION

In the proposal for the Phase II contract, we suggested further investigation of the ideas described in Chapter 8 of the book *Mathematics of Data Fusion*. Specifically, we proposed to further examine (1) components of information, (2) subjective components of information, and so on. A great deal of progress—far in excess of what was originally proposed—has been made, though most of it has been accomplished under two consecutive Air Force Research Laboratory contracts (sections C.3, C.8). Under this work, information-based MoEs have been implemented and used to test a multi-hypothesis correlator-tracker-classifier algorithm in simple two-dimensional scenarios. (Simple scenarios are necessary for such studies: when scenarios become too complex and one observes misbehavior, it becomes very difficult to decide whether the misbehavior is due to the MoEs or to the algorithm being measured.) This work is described more fully in references **[15,17,33,140,141]**. See also **[139]** for an early approach devised by LMTS.

Under the Phase II USARO contract, work was directed at the concept of "multitarget miss distance." FISST provides a natural concept of distance in multitarget problems. In single-target problems, the miss distance between an estimated track $\mathbf{x}$ and a ground truth track $\mathbf{g}$ is the Euclidean distance $\|\mathbf{x} - \mathbf{g}\|$. Recent suggestions have been made to construct multitarget miss-distance MoEs by using some optimal assignment algorithm to associate estimated tracks with ground truth tracks, and then compute the average Euclidean miss distance between those tracks that have been deemed to be associated.

FISST, by way of contrast, provides a *natural* concept of multitarget miss distance called the *Hausdorff distance*. It does not rely on optimal assignment—or, in particular, on any specific optimal association algorithm. It is "natural" in the sense that it metrizes the Mathéron topology on the space of finite subsets. The Hausdorff distance is well-known in image signal processing, and efficient algorithms exist for its computation. It is defined by

$$d_{Haus}(G,X) = \max\{d_0(G,X),d_0(X,G)\}$$

$$d_0(G,X) = \max_{\mathbf{g} \in G} \min_{\mathbf{x} \in X} \|\mathbf{g} - \mathbf{x}\|$$

16

It provides a "worst case" definition of multitarget miss distance. For example, the distance between G = {g_1 ,g_2} and X = {x_1 ,x_2} can be shown to be

$$d_{Haus}(\{g_1 ,g_2\}, \{x_1 ,x_2\}) = \min\{ \max\{\|g_1 - x_1\|, \|g_2 - x_2\|\} , \max\{\|g_1 - x_2\|, \|g_2 - x_1\|\} \}$$

whereas the distance between between G = {g_1 ,g_2} and X = {x} is

$$d_{Haus}(\{g_1 ,g_2\}, \{x\}) = \max\{\|g_1 - x\|, \|g_2 - x\|\}$$

This distance concept has been implemented and tested in the AFRL contracts. It appears to be very promising. See references [33,141].

B.3 PROGRESS IN MEASUREMENT MODELS FOR AMBIGUOUS EVIDENCE

In the proposal for the Phase II contract, we suggested further investigation of: (1) recursive tracking and ID of stationary and moving targets using input data which is precise (e.g. complete radar returns), precise but incomplete (e.g., bearings-only ESM sensors), imprecise (e.g., ambiguous attributes in High Range Resolution radar), vague (i.e., natural-language reports), and contingent (i.e., rules). During the Phase II contract, we showed how to construct *generalized likelihood functions* for ambiguous data and use them in recursive, Bayes-rule filtering algorithms. This work is described more fully in [14,73,92], [62, pp. 63-66], and [60, pp. 14-27 to 14-29] We also initiated an analysis of methods for dealing with problems in which data is exact, but the associated likelihood function cannot be characterized with certainty.

B.3.1 Generalized Likelihood Functions for Ambiguous Data.
Given a random set model of an ambiguous observation (see section A.2.5), the next step in a strict Bayesian formulation of the ambiguous-data problem would be to specify a *likelihood function for ambiguous evidence* that models our understanding of how likely it is that we will observe the specific ambiguous datum Θ, given that a target of state x is present. At this point, however, we immediately encounter practical problems. The required likelihood function must have the form

$$f(\Theta \mid x) = \frac{\Pr(\Re = \Theta, X = x)}{\Pr(X = x)}$$

where $\Re$ is a random variable that ranges over all random closed subsets Θ of measurement space. (These are actually zero-valued probabilities; we are using discrete-variable notation to keep the discussion simple.) However, $f(\Theta|x)$ cannot be a likelihood function unless it satisfies a normality equation $\int f(\Theta|x)d\Theta = 1$ where $\int \cdot d\Theta$ is an integral that sums over all closed random subsets of measurement space. It is very unclear how one would go about constructing a likelihood function $f(\Theta|x)$ that not only models a particular real-world situation but, also, provably integrates to unity. If we knew enough to specify $f(\Theta|x)$ with such exactitude, it would probably also be possible to construct a high-fidelity *conventional* likelihood $f(z|x)$. To address this problem, FISST employs an engineering compromise based on the fact that Bayes' rule is very general: It applies to all events and not just those having the specific Bayesian form E_Θ = "$\Re = \Theta$". That is, Bayes' rule states that $\Pr(E_1| E_2) \Pr (E_2) = \Pr(E_2| E_1) \Pr(E_1)$ for any events E_1, E_2. Consequently let E_Θ be any event with some specified functional dependence on the ambiguous measurement Θ—for example, E_Θ = "$\Theta \supseteq \Xi$" or E_Θ = "$\Theta \cap \Xi \neq \varnothing$" where Θ , Ξ are random closed subsets of observation space. Then

$$f(x \mid E_\Theta) = \frac{\Pr(X = x, E_\Theta)}{\Pr(E_\Theta)} = \frac{\rho(\Theta \mid x)f_0(x)}{\Pr(E_\Theta)}$$

where $f_0(x) = f_X(x) = \Pr(X=x)$ is the prior distribution on x and where $\rho(\Theta|x) = \Pr(E_\Theta|X=x)$ is called a *generalized likelihood function*. Notice that $\rho(\Theta|x)$ will almost always be *unnormalized* (i.e., $\int$

17

$f(\Theta|\mathbf{x})d\Theta \neq 1$) since events E_Θ are not mutually exclusive. Joint generalized likelihood functions can be defined in the same way. Given this, Bayes' rule can be used to compute posterior densities conditioned on ambiguous data modeled by closed random subsets $\Theta_1,\ldots,\Theta_m$ in the usual way:

$$f(\mathbf{x}\mid\Theta_1,\ldots,\Theta_m) \propto \rho(\Theta_1,\ldots,\Theta_m\mid\mathbf{x})f_0(\mathbf{x})$$

with normalization constant $\rho(\Theta_1,\ldots,\Theta_m) = \int \rho(\Theta_1,\ldots,\Theta_m|\mathbf{x})f_0(\mathbf{x})d\mathbf{x}$.

As a simple example [92, pp. 65- 66], assume that both states x and observations z are in the set $\mathbb{R}$ of real numbers. Assume that ambiguous observations have the form $\Theta_{z_0} = \Sigma_A(g_{z_0})$ (see section A.2.5-1) where the fuzzy membership function g_{z_0} on measurement space is:

$$g_{z_0}(z) = \exp\left(-\frac{(z-z_0)^2}{2\sigma_0^2}\right)$$

for all $z \in \mathbb{R}$. To construct a generalized likelihood function for such data, for each target state x we must have an associated "ambiguous signature" of the form $\Xi_x = \Sigma_A(h_x)$, which is our model of what a typical ambiguous observation looks like. Assume that the fuzzy membership function h_x is:

$$h_x(z) = \exp\left(-\frac{(z-z_x)^2}{2\sigma_x^2}\right)$$

Furthermore, we say that an ambiguous observation Θ_z "matches" or "resembles" the ambiguous signature Ξ_x corresponding to the target x if $\Theta_z \cap \Xi_x \neq \varnothing$. (That is, data resembles signature if the two do not contract each other.) Given this, it can be shown that the generalized likelihood function is

$$\rho(\Theta_{z_0}\mid x) = \Pr\left(\Theta_{z_0}\cap\Xi_x \neq \varnothing\right) = \exp\left(-\frac{(z_0-z_x)^2}{2(\sigma_0+\sigma_x)^2}\right)$$

Assume, finally, that $\sigma_x = \sigma$ is constant. Then $\rho(\Theta_{z_0}\mid x) \propto N_{(\sigma+\sigma_0)^2}(z-z_x)$ as a function of x, where $N_{(\sigma+\sigma_0)^2}(z-z_x)$ is the (conventional) likelihood function for the nonlinear measurement model $z = z_x + v$ where v is a zero-mean Gaussian noise process with variance $(\sigma+\sigma_0)^2$. Therefore,

$$f_{k|k}(x\mid\Theta_{z_1},\ldots,\Theta_{z_k}) = f_{k|k}(x\mid z_1,\ldots,z_k)$$

That is: If the fuzzy models h_x have identical Gaussian shapes then a FISST Bayes-rule filter drawing upon fuzzy data behaves exactly like a conventional Bayes nonlinear filter drawing upon ordinary data.

B.3.2 Generalized Likelihood Functions for Imperfectly Characterized Precise Data. A different but related problem occurs when the data z is precise but the corresponding likelihood function $f(z|x)$ is not known with certainty. Under the Phase II contract, LMTS initiated a study of random set-based uncertainty management methods for this class of problems. Much research has been done in "robust estimation," using the techniques first popularized by Huber [35]. In such approaches, one assumes that the likelihood function $L_z(x) = f(z|x)$ is imprecisely known, in the sense that it is known only to belong to some class $\mathfrak{S}$ of density functions. (For example, this class can consist of all functions f that are "close" to some nominal value f_0 , where "close" is defined by some norm on functions: $\|f - f_0\| < \varepsilon$. Another example is the "ε-contamination model," in which the unknown density is assumed to have the form $(1-\varepsilon)f_0 + \varepsilon g$ for g in some class of probability distributions.) Under such assumptions, it is often possible to estimate the unknown state $\mathbf{x}$ robustly in a manner that is optimal in some explicitly specified sense. However, there is a fundamental paradox associated with any approach that is based on a "certain representation of uncertainty." *In such approaches, the uncertainty model is chosen for its mathematical tractability rather than its pertinence to the structure of uncertainty, since this uncertainty is caused by*

ignorance rather than random phenomena. That is: How does one know that the assumed "certain uncertainty model" bears any resemblance to the actual structure of ignorance in the problem?

In our work, LMTS has taken a different point of view: The purpose of an uncertainty model is to *hedge the estimation process against inherently unknowable uncertainties*, rather than to try to optimally estimate using an assumed but possibly irrelevant model of the uncertainty. Briefly, our approach is based on assuming that enough is known about the underlying likelihood function that it can be "trapped" in a *random error bar*: $L_z(\mathbf{x}) \in J_z(\mathbf{x})$, where for each fixed $\mathbf{z}$ and each fixed $\mathbf{x}$, $J_z(\mathbf{x})$ is a random positive interval (i.e., a random closed interval consisting of positive real numbers). That is, the quantity

$$q_j(\mathbf{x}) = \Pr(J_z(\mathbf{x}) = I_j)$$

represents the degree of our belief that the actual value of the likelihood function at $\mathbf{x}$ can be found in the interval I_j. Using this model, *any* nonnegative-valued function $L(\mathbf{x})$ such that $L(\mathbf{x}) \in J_z(\mathbf{x})$ for all $\mathbf{x}$, is a plausible likelihood function for the observation-value $\mathbf{z}$. If we have a sequence of independent, identically distributed observations $\mathbf{z}_1, \ldots, \mathbf{z}_m$, then using *interval arithmetic* we can form the random interval-valued function

$$J_m(\mathbf{x}) = J_{\mathbf{z}_1,\ldots,\mathbf{z}_m}(\mathbf{x}) = J_{\mathbf{z}_1}(\mathbf{x}) \cdots J_{\mathbf{z}_m}(\mathbf{x})$$

where the product of positive intervals is defined by $[a,b]\cdot[c,d] = [ac,bd]$. This function is, in turn, a random error bar for the nominal joint likelihood function:

$$L_{\mathbf{z}_1,\ldots,\mathbf{z}_m}(\mathbf{x}) = L_{\mathbf{z}_1}(\mathbf{x}) \cdots L_{\mathbf{z}_m}(\mathbf{x}) \in J_{\mathbf{z}_1,\ldots,\mathbf{z}_m}(\mathbf{x})$$

Let $\mathcal{L}$ denote the set of all plausible likelihood functions. For each $L \in \mathcal{L}$ we can construct the arg-supremum $\mathbf{x}_L = \text{argsup}_\mathbf{x} L(\mathbf{x})$. This can be a single state-vector, or it can be an infinite subset of such vectors. The *interval-valued argsup* of J_m is the subset of state vectors defined as

$$\text{intargsup}_\mathbf{x} \, J_z(\mathbf{x}) = \{ \, \mathbf{x}_L \mid L \in \mathcal{L} \, \}$$

That is, it is the subset of all *consistent argsup's*. We can derive a specific formula for the interval argsup. For any interval $[a,b]$ define $\overline{[a,b]} = b$ and $\underline{[a,b]} = a$. Then it is easily shown that

$$\text{int}\arg\sup_\mathbf{x} J_m(\mathbf{x}) = \left\{ \mathbf{y} \mid \overline{J_m(\mathbf{y})} \geq \sup_\mathbf{w} J_m(\mathbf{w}) \right\}$$

Furthermore, this definition is compatible with fuzzy-logic representations of uncertainty. That is, suppose that the random interval is the random set associated with a convex fuzzy subset: $J_m(\mathbf{x}) = \Sigma_A(f_\mathbf{x})$. Then it can be shown that the interval argsup is identical to the following "fuzzy argsup":

$$(\arg\sup_\mathbf{x} f_\mathbf{x})(\mathbf{y}) = \Pr\left(\overline{\Sigma_A(f_\mathbf{y})} \geq \underline{\sup_\mathbf{u} \Sigma_A(f_\mathbf{u})}\right)$$

These formulas are being implemented and investigated under another contract (section C.10).

B.4 PROGRESS IN LEVELS 2 AND 3 INFORMATION FUSION

Three years ago, we proposed work to determine whether or not FISST techniques could be extended to Levels 2 and 3 information fusion, i.e. Situation Assessment and Threat Assessment. At that time, our approach to Situation Assessment was to use an idea suggested earlier in the LMTS paper [72, pp. 85-86]. There, we argued that Situation Assessment could be based on multitarget densities of the form $f_{k|k}(X|\mathbf{g})$, which describe how likely it is that a set X of targets would be that consisting of the constituent units of the group-target state $\mathbf{g}$. This general viewpoint proved to be very fruitful and resulted in what LMTS believes is *a genuine conceptual breakthrough in Level 2 information fusion*. Specifically, we have (1) shown how to set up the problem in a correct Bayesian fashion; (2) identified the optimal (but

also computationally intractable) Bayesian solution to the problem; and (3) identified a principled computational approach. Our results can be summarized as follows (see **[56]** for more details).

The objective of Level 2 fusion is to detect, identify, and track not individual targets but rather *group targets* such as infantry battalions, tank columns, artillery chevrons, aircraft sorties, aircraft carrier groups, etc. Level 2 fusion is also often called *force aggregation*. Force aggregation presents a major theoretical and practical challenge. The major reason for this is the *fundamental difficulty involved in deducing the existence and identity of elastically specified, possibly motionless, and possibly physically interleaved group targets using data that is generated not by the groups themselves but rather indirectly by their constituent units*. This sort of complexity means that group-target tracking—the most common type of force aggregation—cannot be viewed in isolation. To be most effective, *the conflicting objectives of group-target tracking, group-target detection, and group-target identification should be optimally integrated.*

That is, a group target cannot be tracked as a group unless we have first decided that it is not just a target group, i.e. some unrelated collection of point targets that happen to be moving together (e.g., a truck caravan and a tank column traveling together on the same road). A group target that is in motion may be detectable because its constituent units have similar velocities. If it is motionless, however, it can be detected only by the presence of characterizing features such as particular geometries (columns, chevrons, wedges, echelons, laagers, etc.), or particular RF transmissions or sequences of such transmissions. Being able to detect and track a group target may be of little tactical interest if we cannot also determine whether or not it is a tank column or a truck column. Two moving but interleaved group targets may be difficult to track unless we can determine that (for example) one is an armored battalion whereas the other is a mobile infantry battalion.

Our solution to the force aggregation is as follows:

(1) recognize that the unknown random Bayesian state-parameter in a force aggregation problem is a special kind of random process called a *cluster process*;
(2) the optimal method for propagating this cluster process through time is a suitable generalization of the multisensor-multitarget Bayes filtering equations of section 1-1-1 of Appendix 1;
(3) though these equations will be computationally intractable in most situations, under high-SNR conditions it may be possible to approximate them using suitable generalizations of the concept of a multitarget first-order moment density (as defined in section B.6.2 below)

The ordinary multisensor-multitarget problem has two "levels": a hidden target-track level (the space of unknown target states); and a visible observation level (the space of known measurements). Underlying everything is an unknown generating "center process" *or mother process*—i.e., the random multitarget track-process $\Xi_{k|k}$ that consists of a randomly-varying number of randomly-varying state-vectors. Each state-vector $\mathbf{x}$ in the track-process is, in turn, associated with ("marked" by) a *daughter process*—i.e., the random observation-process ("cluster") $\Sigma_k^{\mathbf{x}}$ that consists of a randomly-varying number of randomly-varying observations generated by $\mathbf{x}$. Stated somewhat differently, a complete description of the conventional multisensor-multitarget system at any instant is a finite set $\mathbb{Z} = \{(\mathbf{x}_1, Z_1),...,(\mathbf{x}_n, Z_n)\}$ of pairs where Z_j is the observation-set currently generated by the target with state $\mathbf{x}_j$. So, for each instantiation of the mother track-process, $\Xi_{k|k} = X = \{\mathbf{x}_1,...,\mathbf{x}_n\}$, the system with multitarget state X is a random finite set of pairs $\mathcal{Z}_k|_X = \{(\mathbf{x}_1, \Sigma_k^{\mathbf{x}_1}),...,(\mathbf{x}_n, \Sigma_n^{\mathbf{x}_n})\}$. Since in a Bayesian analysis the target state-sets vary randomly as well, the total statistical representation of the multisensor-multitarget system is the *random finite subset* $\mathcal{Z}_k = \cup_{\mathbf{x} \in \Xi_{k|k}} \{(\mathbf{x}, \Sigma_k^{\mathbf{x}})\}$ of the space of pairs. The random process is an example of a *cluster process with center process* $\Xi_{k|k}$. **[13,39]**

20

Unlike a conventional multisensor-multitarget problem, a force aggregation problem has *three* layers: a twice-hidden group-target layer (the space of unknown states of the group targets); a singly-hidden layer (the space of unknown states of the ordinary targets); and the visible observation layer. In this problem, *the system of unknown quantities is itself a cluster process*. Underlying everything else is a mother process $\Gamma_{k|k}$—i.e., the random variable whose instantiations $\Gamma_{k|k} = G$ are finite sets $G = \{g_1,...,g_e\}$ of unknown group-target states g_j. (At it simplest, g_j can belong to a Euclidean vector space, e.g.: $g_j = (x,v,N,\tau,\gamma)$ where x is the geometric centroid and v is its velocity; N is the number of targets; τ is the type; and γ is a geometric-shape parameter such as chevron, column, etc. More generally, g can be a function in some Hilbert space.) Each group-target state g is "marked" by a daughter process $\Xi^g_{k|k}$—i.e., the random variable whose instantiations $\Xi^g_{k|k} = X$ are finite sets $X = \{x_1,...,x_n\}$ of the unknown target-states x that comprise the group-target with state g. Stated in different terms, the *complete state specification of a group target* is a finite set set $\mathbb{X} = \{(g_1, X_1),...,(g_e, X_e)\}$ where X_j is the set of the states of the individual targets that constitute the group target g_j. So, for each instantiation $\Gamma_{k|k} = G = \{g_1,...,g_e\}$ of the mother group-track process, the system with multigroup state G is a random finite set of pairs pairs $\mathcal{X}_{k|k}|_G = \{(g_1,\Sigma^{g_1}_{k|k}),...,(g_e,\Sigma^{g_e}_{k|k})\}$, where $\Xi^g_{k|k}$ is the daughter track-process generated by the group target with state g. Since in a Bayesian analysis the multigroup target state-sets vary randomly as well as the observation-set, the total statistical representation of the multisensor-multigroup system is the random finite *subset* $\mathcal{X}_{k|k} = \cup_{g \in \Gamma_{k|k}} \{(g,\Xi^g_{k|k})\}$ of the space of pairs.

Given this, the *group multitarget density function*

$$f_{k|k}(X \mid g) = \frac{\delta}{\delta X} \beta_{\Xi^g_{k|k}}(S) = \left[\frac{\delta}{\delta X}\Pr\left(\Xi^g_{k|k} \subseteq S\right)\right]_{S=\varnothing} = \Pr\left(\Xi^g_{k|k} = S\right)$$

is the multitarget density function of the daughter track-process $\Xi^g_{k|k}$, as described in our Phase II proposal three years ago. It provides a probabilistic definition of any specific group target g. For example, if g is a tank chevron with specified orientation and nominal location then $f_{k|k}(X|g)$ will be small if X is not a tank chevron with the specified orientation and location. On the other hand, $f_{k|k}(X|g)$ will be large if X resembles a tank chevron with the specified orientation and location. The more X resembles a group target g, the greater the value of $f_{k|k}(X|g)$. The quantity $f_{k|k}(X|g)$ also models various kinds of ambiguity—e.g., the fact that a battalion can have varying numbers of platforms of a given type and nevertheless still be a battalion. The more ambiguous the definition of a group target g, the more concentrated $f_{k|k}(X|g)$ will be around some specific group X of targets.

We are now in a position to describe the optimal solution to the Level 2 information fusion problem. Let $\mathbb{X} = \{(g_1, X_1),...,(g_e, X_e)\}$ be a collection of group targets $g_1,...,g_e$ with respective target-sets $X_1,...,X_e$, and let $f_{k|k}(\mathbb{X}|Z^{(k)})$ be the posterior density on $\mathbb{X}$. Intuitively speaking, this distribution is

$$f_{k|k}(\mathbb{X}|Z^{(k)}) = \Pr\left(\Gamma_{k|k} = \{g_1,...,g_e\}, \Xi^{g_1}_{k|k} = X_1,...,\Xi^{g_e}_{k|k} = X_e\right)$$

Our goal is to propagate $f_{k|k}(\mathbb{X}|Z^{(k)})$ through time and, at each time-step, to estimate the multigroup state $\{(\hat{g}_1,\hat{X}_1),...,(\hat{g}_{\hat{e}},\hat{X}_{\hat{e}})\}$ that best explains the data. If we are successful, then at any time-step k we would have a simultaneous, joint estimate of the complete state of the force-aggregation problem: a collection $\{(\hat{g}_1,...,\hat{g}_{\hat{e}}\}$ of estimated group targets and their number, together with a collection $\{\hat{X}_1,...,\hat{X}_{\hat{e}}\}$ of their respective track-sets. Optimal propagation of the multisensor-multigroup posterior is accomplished using the following group-target analog of the multisensor-mulititarget Bayes filter (Equation 3 of section 1-1-1 of Appendix 1):

$$f_{k+1|k}(\mathbb{X}|Z^{(k)}) = \int f_{k+1|k}(\mathbb{X}|\mathbb{W}) \, f_{k|k}(\mathbb{W}|Z^{(k)}) \, \delta\mathbb{W}$$

$$f_{k+1|k+1}(\mathbb{X}|Z^{(k+1)}) \propto f_{k+1}(Z_{k+1}|\mathbb{X}) \, f_{k+1|k}(\mathbb{X}|Z^{(k)})$$

where: $\mathbb{X} = \{(\mathbf{g}_1, X_1),...,(\mathbf{g}_e, X_e)\}$ is the unknown multigroup state-set; $Z^{(k)} = \{Z_1,...,Z_k\}$ is the time-series of collected observation-sets at time-step k; $f_k(Z|\mathbb{X})$ is the multisensor-multitarget likelihood function; $f_{k+1|k}(\mathbb{X}|\mathbb{W})$ is the multitarget Markov transition density; $f_{k|k}(\mathbb{X}|Z^{(k)})$ is the multitarget posterior distribution at time-step k; $f_{k+1|k}(\mathbb{X}|Z^{(k)})$ is the prediction of this posterior to time-step $k+1$; and where

$$f_{k+1}(Z_{k+1}|Z^{(k)}) = \int f_{k+1|k}(Z_{k+1}|\mathbb{X}) \, f_{k|k}(\mathbb{X}|Z^{(k)}) \, \delta\mathbb{X}$$

is the Bayes normalization constant.

Just as the conventional multisensor-multitarget Bayes filter cannot be copied blindly from the single-sensor, single-target Bayes filter, so the Bayes multisensor-multigroup filter cannot be copied blindly from the multisensor-multigroup filter. The first and most obvious reason is that the integrals $\int \cdot \, \delta\mathbb{X}$ occurring in the multigroup filter equations are much more complex than ordinary set integrals. Rather they are *generalized set integrals*—what we call *group integrals*—that sum over all finite sets of group targets $\mathbb{X} = \{(\mathbf{g}_1, X_1),...,(\mathbf{g}_e, X_e)\}$:

$$\int f(\mathbb{X})\delta\mathbb{X} = \sum_{j=0}^{\infty} \frac{1}{j!} \int f(\{(\mathbf{g}_1, X_1),...,(\mathbf{g}_j, X_j)\}) d\mathbf{g}_1 ... d\mathbf{g}_j \delta X_1 \cdots \delta X_j$$

where each of the indicated integrals $\int \cdot \, \delta\mathbb{X}_j$ is an ordinary set integral. The second reason is that we cannot merely assume the existence of the multisensor-multigroup likelihood function $f_k(Z|\mathbb{X})$ and the multigroup Markov density $f_{k+1|k}(\mathbb{Y}|\mathbb{X})$, but rather must *construct* them. This requires a *generalized set derivative* (which we call a *group derivative*), which we will not describe here. A final reason why the multisensor-multigroup filter cannot be copied blindly is the fact that the naïve maximum *a posteriori* (MAP) estimate requires even more care with multigroup states than with multitarget states.

A final issue is computability. Clearly, if the multisensor-multitarget Bayes filter is computationally intractable in most circumstances, then the multisensor-multigroup Bayes filter will almost always be so. Consequently, it is of mere mathematical interest without the existence of drastic but principled ways of implementing it. Such a method was devised under the Phase II contract—a filter based on the concept of a multitarget first-order moment density—and will be described in section B.6.2 below.

B.5 PROGRESS IN TRACK-TO-TRACK FUSION

Three years ago, we proposed to investigate two different approaches to track-to-track fusion: an optimal approach, assuming that double-counted observations between two sources are known *a priori*; and a robust approach, assuming that nothing whatsoever is known about correlations between the sources. We succeeded in both of these efforts, which are described in [63].

B.5.1 Optimal Track-to-Track Fusion. Our proposed approach was to use the almost-parallel worlds principle (APWOP) to generalize to the multitarget case an optimal multi-source, single-target approach devised in 1990 by Chong, Mori, and Chang. Assume that, at time-step k, two multitarget fusion algorithms generate "local" multitarget posteriors $f^{[s]}_{k|k}(X|Z^{(k)})$ for $s = 1,2$. These algorithms share reports from some of the same sensors, which means that a central fusion site must construct its own multitarget posterior $f_{k|k}(X|Z^{(k)})$, taking *double-counting* into account. Let $Z^{[s]}_k$ denote the multisensor-multitarget observation-set collected at time-step k by the s'th algorithm, and let $Z_{[s]}^{(k)}$ be the time-

sequence of such observations collected at time-step k by the same algorithm. Then we demonstrated that the optimal fusion equation is:

$$f_{k+1|k+1}(X \mid Z^{(k+1)}) \propto \frac{f^{[1]}_{k+1|k+1}(X \mid Z^{(k+1)}_{[1]})}{f^{[1]}_{k+1|k+1}(X \mid Z^{(k+1)}_{[1]} \cap Z^{(k+1)}_{[2]}, Z^{(k+1)}_{[1]})} \cdot \frac{f^{[2]}_{k+1|k+1}(X \mid Z^{(k+1)}_{[2]})}{f^{[2]}_{k+1|k}(X \mid Z^{(k)}_{[2]})} \cdot f_{k+1|k}(X \mid Z^{(k)})$$

where $Z^{(k+1)}_{[1]} \cap Z^{(k+1)}_{[2]}$ is the set of double-counted observations.

B.5.2 Robust Track-to-Track Fusion. Our proposed approach was to use the APWOP to generalize to the multitarget case the "covariance intersection (CI)" approach introduced by Uhlmann and Julier in the mid-1990s. CI is a method for fusing two or more Gaussian sources (track with track, track with report, report with report) that protects against *worst-case correlations* between the sources. Our approach was to first generalize CI to the arbitrary (i.e., non-Gaussian) single-target case, and then use the APWOP to generalize further to the multitarget case. Let $f_0(X|Z_0^{(k)})$ and $f_1(X|Z_1^{(k)})$ be the multitarget posteriors produced by two fusion algorithms, drawing respectively on their own streams $Z_0^{(k)}$ and $Z_1^{(k)}$ of multisensor-multitarget observation-sets. Define the quantity

$$s(\omega) = \sup_X \frac{c^{|X|}}{|X|!} f_\omega(X \mid Z_0^{(k)}, Z_1^{(k)})$$

where c is a suitable constant and where

$$f_\omega(X \mid Z_0^{(k)}, Z_1^{(k)}) = \frac{f_0(X \mid Z_0^{(k)})^{1-\omega} f_1(X \mid Z_1^{(k)})^\omega}{\int f_0(Y \mid Z_0^{(k)})^{1-\omega} f_1(Y \mid Z_1^{(k)})^\omega \delta Y}$$

Let $a = argsup_\omega \ s(\omega)$. Then we showed that the generalization of CI to the multitarget track-to-track fusion problem is the fused multitarget density $f(X|Z_0^{(k)}, Z_1^{(k)}) = f_0(X|Z_0^{(k)}, Z_1^{(k)})$. We similarly showed how to robustly fuse a multisource-multitarget report with a multitarget track and with another multisource-multitarget report.

B.6 PROGRESS IN COMPUTATIONAL TECHNIQUES

Though the multisensor-multitarget Bayes filter (Equation 3 of section 1-1-1 of Appendix 1) is optimal, it will be computationally intractable in most situations. Consequently, multitarget nonlinear filtering will be of little practical interest in real-time problems unless drastic but principled approximation strategies can be devised. Three years ago in our Phase II proposal, we suggested the following lines of attack: (1) approximate computation of permanents of matrices; (2) approximation by Gaussian sums; (3) asymptotic approximations of integrals; and (4) computational statistical mechanics. Unfortunately, analysis conducted during the project indicates that none of these approaches are likely to be feasible in real-time operation.

In their place, we devised two new approaches that appear to be much more promising. These are approximation based on a multitarget generalization of: (1) the Gaussian density, called a "para-Gaussian"; and (2) first-order statistical moments. In both cases, the approximation technique is based on an analogy with the Kalman filter. While the first method has not as yet been implemented, the second method is being investigated in other R&D contracts (sections C.6, C.7, C.11). Finally, it is also worth mentioning that LMTS is investigating a third approach under internal R&D funding: multitarget filtering based on particle system filters. This method is also reported below. At this time, our preliminary assessment of these techniques is as follows:

(1) *Particle-systems approximation*: Very flexible; appropriate for near-exact real-time implementation, for scenarios involving up to four or five targets in 2-D scenarios (fewer in 3-D ones).

23

(2) *Para-Gaussian approximation:* Regime of appropriateness unknown; each choice of maximum number of targets requires implementation of a different algorithm.

(3) *First-order statistical moment approximation:* Flexible; appropriate for real-time approximate implementation of the multitarget Bayes filter in high-density scenarios, assuming very high SNR (if high localization accuracy is required).

B.6.1 Multitarget Filtering Based on a "Para-Gaussian" Approximation. These results have been reported in [64] and [62, pp. 49-52]. This approximation method for multitarget filtering is based on the following *direct* analogy with the familiar Gaussian approximation. There are three major sources of computational load in the single-sensor, single-target nonlinear filtering equations (Equation 1 and Equation 2 of section 1-1-1 of Appendix 1). First, two numerical integrations are required to compute the prediction integral and the Bayes normalization constant; and second, computations involving the indefinitely large number of parameters that are required to specify the evolving posterior distribution. The conventional Gaussian approximation addresses both concerns as follows. In statistical theory, a probability density $f_0(\mathbf{x})$ is said to belong to a family of *conjugate priors* for a given likelihood function $f(\mathbf{z}|\mathbf{x})$ if the posterior distribution $f(\mathbf{x}|\mathbf{z}) \propto f(\mathbf{z}|\mathbf{x}) f_0(\mathbf{x})$ belongs to this same family [52, pp. 59-65] If such a family exists then computation using Bayes' rule can often be greatly simplified. In particular, if $f(\mathbf{z}|\mathbf{x})$ is Gaussian then the identity

$$N_A(\mathbf{x}\text{-}\mathbf{a}) \, N_B(\mathbf{x}\text{-}\mathbf{b}) = N_{A+B}(\mathbf{a}\text{-}\mathbf{b}) \, N_C(\mathbf{x}\text{-}\mathbf{c})$$

(where $C^{-1} = A^{-1} + B^{-1}$ and $C^{-1}\mathbf{c} = A^{-1}\mathbf{a} + B^{-1}\mathbf{b}$) shows that the corresponding family of conjugate priors is just the family of Gaussian distributions. Moreover, the same equation shows that integrals of products of Gaussians—in particular the prediction integral and the Bayes normalization constant—satisfy the following *closed-form integrability property*:

$$\int N_A(\mathbf{x}\text{-}\mathbf{a}) \, N_B(\mathbf{x}\text{-}\mathbf{b})d\mathbf{x} = N_{A+B}(\mathbf{a}\text{-}\mathbf{b})$$

Suppose now that $f(Z|X)$ is the multitarget likelihood for a single Gaussian sensor, with both missed detections and false alarms being taken into account. From a computational point of view life would be greatly simplified if, as in the single-target case, we could find a family of *multitarget priors* $f_0(X)$ that are conjugate to this likelihood and that have closed-form set-integrability properties. Unfortunately, it appears that no such family exists. However, there *is* a family of multitarget distributions—the "para-Gaussian" multitarget distributions—that is not conjugate but *does* have a closed-form set-integrability property. If we use this family, computational tractability becomes potentially feasible even if we use motion models that do not assume that the number of targets is fixed.

Specifically, suppose that we have a single Gaussian sensor with missed detections. From the FISST multitarget calculus we know [24, pp. 166-168] that the multitarget likelihood is

$$f_0(Z \mid X) = p_D^m (1 - p_D)^{n-m} \sum_{1 \le i_1 \ne \ldots \ne i_m \le n} N_Q(\mathbf{z}_1 - B\mathbf{x}_{i_1}) \cdots N_Q(\mathbf{z}_m - B\mathbf{x}_{i_m})$$

for $Z = \{\mathbf{z}_1, \ldots, \mathbf{z}_m\}$ and for $X = \{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$. If in addition the Gaussian sensor is corrupted by a statistically independent, state-independent clutter process with density $\kappa(Z)$ then we also know that the multitarget likelihood is

$$f_{targets + clutter}(Z|X) = \sum_{W \subseteq Z} f_0(W|X) \, \kappa(Z\text{-}W)$$

So, for $Y = \{\mathbf{y}_1, \ldots, \mathbf{y}_r\}$ with $r \ge n$ define the *para-Gaussian multitarget density* $N_{Q,q,\kappa}(X|Y)$ by

24

$$N_{Q,q}(X \mid Y) = \frac{q(n \mid r)}{C_{r,n}} \sum_{1 \leq i_1 \neq \ldots \neq i_n \leq r} N_Q(\mathbf{x}_1 - \mathbf{y}_{i_1}) \cdots N_Q(\mathbf{x}_m - \mathbf{y}_{i_m})$$

$$N_{Q,q,\kappa}(X \mid Y) = \sum_{W \subseteq X} N_{Q,q}(X \mid Y) \kappa(X - W)$$

where $q(n|r) \geq 0$ for all j, where $q(n|r) = 0$ if $n > r$ or $n < 0$, and where $\sum_n^r q(n \mid r) = 1$. During the Phase I contract we showed that para-Gaussians obey the following closed-form set-integrability property **[24, pp. 243-244]**:

$$\int N_{P,p,\kappa}(Z|X) \, N_{Q,q}(Z|X) \delta X = N_{P+Q, \, p \otimes q, \kappa}(Z|X)$$

where $(p \otimes q)(k \mid i) = \Sigma_{j=k}^i p(k \mid j) q(j \mid i)$.

The computational advantage resulting from this fact suggests the following multitarget analog of the Gaussian approximation. Assume that the underlying sensor is Gaussian with a statistically independent, state-independent clutter process κ. Then the multitarget likelihood function of the sensor has the para-Gaussian form $f_k(Z \mid X) = N_{Q_k, q_k \kappa_k}(Z \mid X)$. Let $\mathbf{x}_{k+1} = \Phi_k(\mathbf{x}_k)$ be the deterministic motion-update at time-step $k+1$ and define $\Phi_k(X) = \{\Phi_k(\mathbf{x}_1), \ldots, \Phi_k(\mathbf{x}_n)\}$. Assume that the multitarget motion model and all multitarget posteriors are para-Gaussian:

$$f_{k+1|k}(Y \mid X) = N_{R_k, r_k}(Y \mid \Phi X)$$

$$f_{k|k}(X \mid Z^{(k)}) = N_{P_k, p_k}(X \mid \hat{X}_k)$$

In other words: (1) target motions are independent; (2) targets can disappear but not appear; and (3) any multitarget posterior can be described by the finite set of parameters P_k, p_k, and $\hat{X}_k$. (This is a fairly drastic simplification since it means that a *single* covariance matrix P_k is forced to describe the uncertainty in the state estimate of *every* target. However, optimal-Bayes multitarget filtering techniques are necessary only when targets are relatively close together—otherwise, the problem can be split up into parallel single-target filters. The uncertainties of tracks that are close together are more likely to be similar than those of targets that are far apart.)

Let $C_{k+1} = R_k + \Phi_k^T P_k \Phi_k$, $c_k = r_k \otimes p_k$, and let Z_{k+1} be the observation-set collected at time-step k. Then the para-Gaussian approximation is based on the following recursion:

- Compute the multitarget prediction integral in closed form:
$$f_{k+1|k}(X|Z^{(k)}) = \int f_{k+1|k}(X|W) f_{k|k}(W|Z^{(k)}) \delta W = N_{C_k, c_k}(X \mid \Phi \hat{X}_k)$$

- Compute the multitarget Bayes normalization constant in closed form:
$$f_k(Z_{k+1}|Z^{(k)}) = \int f_{k+1}(Z_{k+1}|W) f_{k+1|k}(W|Z^{(k)}) \delta W = N_{Q_k + C_{k+1}, q_k \otimes c_{k+1}, \kappa_k}(Z_{k+1} \mid \Phi \hat{X}_k)$$

- Construct the Joint Multitarget Estimate (JoME) of section 1-2 of Appendix 1:
$$\hat{X}_{k+1} = \hat{X}_{k+1|k+1}^{JoME}$$

- Compute the following integrals in closed form (which can also be done):
$$I_n = f_{k+1}(Z_{k+1}|W)^{-1} \int f_{k+1}(Z_{k+1}|\{\mathbf{x}_1, \ldots, \mathbf{x}_n\}) f_{k+1|k}(\{\mathbf{x}_1, \ldots, \mathbf{x}_n\}|Z^{(k)}) d\mathbf{x}_1 \cdots d\mathbf{x}_n$$
- Define $p_{k+1}(n|r) = I_n$ if $0 \leq n \leq r$ and $p_{k+1}(n|r) = 0$ otherwise, as well as $P_{k+1} = (Q_k^{-1} + C_{k+1}^{-1})^{-1}$.

- Return to the first step.

B.6.2 Multitarget Filtering Based on a Multitarget First-Order Moment Approximation. These results have been reported in [16,58,59,65]. This approximation method for multitarget filtering is based on a second, *statistical* anlogy with the Gaussian approximation. In the single-target case, a historically important strategy for side-stepping the computational complexity of the single-sensor, single-target Bayes filtering equations (Equation 1 and Equation 2 of section 1-1-1 of Appendix 1) has been to assume that signal-to-noise ratio is high enough that the first-moment vector and second-moment matrix

$$\mathbf{x}_{k|k} = \int \mathbf{x} f_{k|k}(\mathbf{x}|Z^k)dx, \qquad M_{k|k} = \int \mathbf{x}\mathbf{x}^T f_{k|k}(\mathbf{x}|Z^k)dx$$

are approximate sufficient statistics: $f_k(\mathbf{x}|Z^k) \cong f_{k|k}(\mathbf{x}|\mathbf{x}_{k|k},M_{k|k}) = N_P(\mathbf{x}\text{-}\mathbf{x}_{k|k})$ where $N_P(\mathbf{x}\text{-}\mathbf{x}_{k|k})$ is a Gaussian distribution with covariance matrix $P = M_{k|k} - \mathbf{x}_{k|k}\mathbf{x}_{k|k}^T$. In this case we can propagate $\mathbf{x}_{k|k}$ and $M_{k|k}$ instead of the full posterior distribution $f_{k|k}(\mathbf{x}|Z^k)$ using a Kalman filter. If SNR is so high that the second-order moment can be neglected as well, then $f_{k|k}(\mathbf{x}|Z^k) \cong f_{k|k}(\mathbf{x}|\mathbf{x}_{k|k})$ and we can propagate $\mathbf{x}_{k|k}$ alone using a constant-gain Kalman filter—e.g., the α-β-γ filter.

During the Phase II contract, LMTS demonstrated that that this basic reasoning can be extended to multisensor-multitarget problems—though not in a naïve manner. Whenever one mentions the concept of a "multitarget first-order moment" of a random track-set $\Xi_{k|k}$, what engineers usually expect to see is a *track-valued expectation*—that is, *a set of specific tracks* of the form $E[\Xi_{k|k}] = \{\mathbf{x}_1,...,\mathbf{x}_n\}$ where $\mathbf{x}_1,...,\mathbf{x}_n$ are the tracks in the expectation. Although we attempted to find a theoretically acceptable definition of a track-valued expectation during the Phase II contract [71], we succeeded in doing so only in special cases. Instead, we took a different approach: that of defining a multitarget expectation *indirectly*. That is, one constructs multitarget moments by first specifying some function ϕ that transforms multitarget state-sets $X = \{\mathbf{x}_1,...,\mathbf{x}_n\}$ into elements $\phi(X)$ of some suitably well-behaved vector space. The transformation ϕ should be one-to-one and it should transform set-theoretic operations into corresponding vector-algebra operations—for example, $\phi(X \cup Y) = \phi(X) + \phi(Y)$ whenever $X \cap Y = \varnothing$. In this case we can compute first-order moments of the form $E[\phi(\Xi_{k|k})]$ that will themselves be elements of this vector space. Two obvious candidates, identified by LMTS during the Phase I contract, are the following [24, p. 179]:

$$\delta_X(\mathbf{x}) = \delta_{\mathbf{x}_1}(\mathbf{x}) + ... + \delta_{\mathbf{x}_n}(\mathbf{x})$$

$$\Delta_X(S) = \Delta_{\mathbf{x}_1}(S) + ... + \Delta_{\mathbf{x}_n}(S)$$

where $\delta_w(\mathbf{x})$ denotes the Dirac delta function concentrated at w and where $\Delta_w(\mathbf{x})$ denotes its corresponding Dirac measure: $\Delta_w(S) = 1$ if $\mathbf{x} \in S$ and $\Delta_w(S) = 0$ otherwise.

All of the three following items are interchangeably known as a (simple*) multi-dimensional point process* [131, pp. 100-102], [4,13,115]: the random subset $\Xi_{k|k}$, the random "counting measure" $\Delta_{\Xi_{k|k}}(S)$, and its corresponding random density function $\delta_{\Xi_{k|k}}(\mathbf{x})$.

Given this, the indirect multitarget expectation of a random track-set $\Xi_{k|k}$, which we call the *probability hypothesis density* or PHD, is just the expectation

$$D_{k|k}(\mathbf{x}|Z^{(k)}) = E[\delta_{\Xi_{k|k}}] = \int \delta_X(\mathbf{x}) \cdot f_{k|k}(X|Z^{(k)})\delta X = \int_{X \supseteq \{\mathbf{x}\}} f_{k|k}(X|Z^{(k)})\delta X$$

of the random function $\delta_\Xi(\mathbf{x})$. The PHD concept was first introduced into the information fusion community in 1993 by M.C. Stein and C.L. Winter, as a force aggregation approach [128]. Intuitively speaking, just as the value of the probability density function $f_{k|k}(\mathbf{x}|Z^k)$ of a continuous random vector $\mathbf{X}_{k|k}$ provides a means of describing the zero-probability event $\Pr(\mathbf{X}_{k|k} = \mathbf{x})$, so the PHD $D_{k|k}(\mathbf{x}|Z^{(k)})$ of a finite random track-set $\Xi_{k|k}$ provides a means of describing the zero-probability event $\Pr(\mathbf{x} \in \Xi_{k|k})$. Consequently, $D_{k|k}(\mathbf{x}|Z^{(k)})$ will tend to have maxima approximately at the locations of the targets. (The

PHD is also well-known in point process theory, where its corresponding measure $\mu_{\Xi_{k|k}}(S) = E[\Delta_{\Xi_{k|k}}(S)]$ is called the *first moment measure* [13]; see sections B.6.2 and B.8 below.)

During the Phase II contract LMTS showed that, under the same high-SNR assumption that is necessary for constant-gain Kalman filters in the single-sensor, single-target case, it is possible to derive recursive Bayes filter equations for the PHD. These equations are general enough to include models for disappearance and appearance of targets. Specifically, between measurement collection times, the PHD can be propagated from time-instant k to time-instant $k+1$ by using the prediction integral

$$D_{k+1|k}(\mathbf{y} \mid Z^{(k)}) = \int \left(d_{k+1|k}(\mathbf{x}) f_{k+1|k}(\mathbf{y} \mid \mathbf{x}) + B_{k+1|k}(\mathbf{y} \mid \mathbf{x})\right) D_{k|k}(\mathbf{x} \mid Z^{(k)}) d\mathbf{x}$$

where: (1) $1 - d_{k|k}(\mathbf{x})$ is the probability that a target with state $\mathbf{x}$ at time-step k will disappear from the scene at time-step $k+1$; and (2) $B_{k|k}(\mathbf{y}|\mathbf{x})$ is the PHD of the multitarget density $b_{k|k}(Y|\mathbf{x})$ that describes the likelihood that a target with state $\mathbf{x}$ at time-step k will generate a set X of new targets at time-step $k+1$.

If SNR is high enough, we can derive a similar Bayes-rule update (though it will be approximate rather than exact). In general, the multitarget statistics of the random track-set will be quite complex. If SNR is high enough, however, then we can assume that the track-set obeys simpler, Poisson multitarget statistics. In this case the following approximate equation is the Bayes-update step of the PHD using a new multitarget observation-set Z_{k+1}:

$$D_{k+1|k+1}(\mathbf{x} \mid Z^{(k+1)}) \cong \sum_{z \in Z_{k+1}} \frac{p_D D_{k+1}(\mathbf{z})}{\lambda_{k+1} c_{k+1}(\mathbf{z}) + p_D D_{k+1}(\mathbf{z})} D_{k+1|k+1}(\mathbf{x} \mid \mathbf{z}, Z^{(k)}) + (1 - p_D) D_{k+1|k}(\mathbf{x} \mid Z^{(k)})$$

where: (3) $N_{k+1|k} = \int D_{k+1|k}(\mathbf{x}|Z^{(k)}) d\mathbf{x}$ is the predicted expected number of targets at time-step $k+1$; (4) p_D is the (state-independent) probability of detection of the sensor, assumed to be large enough that $p_D > 1 - N_{k+1|k}^{-1}$ for any k; (5) λ_{k+1} is the average number of Poisson false alarms per data-scan, and $c_{k+1}(\mathbf{z})$ is the (state-independent) distribution of each of these false alarms; (6) $D_{k+1}(\mathbf{z}) = \int f(\mathbf{z}|\mathbf{x}) D_{k+1|k}(\mathbf{x}|Z^{(k)}) d\mathbf{x}$; and (7) the quantity

$$D_{k+1|k+1}(\mathbf{x} \mid \mathbf{z}, Z^{(k)}) = \frac{f(\mathbf{z} \mid \mathbf{x}) D_{k+1|k}(\mathbf{x} \mid Z^{(k)})}{D_{k+1}(\mathbf{z})}$$

is a Bayes-rule-like update of $D_{k+1|k}(\mathbf{x}|Z^{(k)})$ using the observation $\mathbf{z}$. LMTS has also shown that these equations are easily extended to deal with multiple sensors, assuming conditional independence of their observation-sets.

What the above "PHD filtering equations" tell us is that if signal-to-noise ratio is large enough then multitarget detection, tracking, and identification can be accomplished using a process that strongly resembles single-target nonlinear filtering. This being the case, *any* computational nonlinear filtering approach can, in principle, be used to implement these equations. Furthermore, these equations *do not require report-to-track association*. Rather, data association is essentially replaced by *multi-peak extraction*. At each stage, the PHD filter propagates not only the PHD $D_{k|k}(\mathbf{x}|Z^{(k)})$ but also the expected number of targets $N_{k|k} = \int D_{k|k}(\mathbf{x}|Z^{(k)}) d\mathbf{x}$. Consequently, estimation of the multitarget state is accomplished by computing the nearest integer $[N_{k|k}]$ in $N_{k|k}$ and then searching for the $[N_{k|k}]$ largest peaks of $D_{k|k}(\mathbf{x}|Z^{(k)})$. Also, because the PHD filtering equations have the same general form as the conventional recursive Bayes filter (equation 1 of section 1-1-1 of Appendix 1) this means that, in principle, the PHD filter can be implemented using *any* computational nonlinear filtering technique. (See [133] for a related approach.)

Also, it can be shown that PHD filters can be devised for multigroup target scenarios of the kind described in section B.4. In this case, one notices that the multigroup process $\mathcal{X}_{k|k} = \cup_{g \in \Gamma_{k|k}} \{(g, \Xi^g_{k|k})\}$ is equivalent to the random finite track-set

$$\tilde{\Xi}_{k|k} = \cup_{g \in \Gamma_{k|k}} (\{g\} \times \Xi^g_{k|k})$$

Since this is a finite random subset it has a corresponding PHD, which in turn can be propagated through time using slightly modifed versions of the PHD filtering equations.

B.6.3 Multitarget Filtering Based on Particle-Systems Approximation. This approach has been reported in [3]. The single-sensor, single-target recursive Bayes filtering equations (Equation 1 and Equation 2 of section 1-1-1 of Appendix 1) is of only mathematical interest unless it can be implemented for real-time application. A promising implementation approach that has been attracting much attention in recent years is known as "particle-systems filtering." In this approach, the posterior distribution $f_{k|k}(\mathbf{x}|Z^k)$ is approximated using a large group of sampling points $\mathbf{x}$ of the distribution—i.e., more points where the posterior is largest and fewer where it is not. These samples are treated as though they were particles moving through state space. In principle, particles corresponding to small posterior values can be eliminated, and those corresponding to large posterior values can be allowed to spawn ("give birth to," "branch") several new particles. Particle-systems filters have very general convergence properties, and can handle tracking situations (e.g. heavy-tailed models, discontinuous models) that other approaches, e.g. based on the Fokker-Planck (forward-Kolmogorov) equation cannot [46]. Under LMTS internal R&D funding, the University of Alberta at Edmonton has been developing novel new branching-particle filter algorithms. Since particle-systems approaches apply to any state space that is Polish and since point processes form a Polish space, particle systems techniques can be directly extended to multitarget problems. Most recently, the University of Alberta has applied the particles systems approach to multitarget problems [3]. LMTS is currently evaluating the usefulness of this approach to computational multitarget nonlinear filtering.

B.7 PROGRESS IN ALGORITHMIC FEASIBILITY ANALYSIS

In the Phase II contract proposal we suggested the following possibilities for implementing algorithms based on FISST ideas: (1) simultaneous determination of the numbers and locations of targets based on precise and ambiguous evidence; (2) computing a multitarget ROC curve for a multitarget decision problem, e.g. detection of a known target in clutter; (3) measuring the amount of information supplied by a simple data fusion algorithm; and (4) a sensor management algorithm.

Once again, our approach during the Phase II contract was to try to win independent funding from other agencies to look at one or more of these problems (see section A.1). Under two consecutive contracts from AFRL/IFEA we have implemented multisensor-multitarget information MoEs for Levels 1,2, and 4 information fusion (see sections C.3 and C.8). This work has been reported in [15,17,140,141]. Under contract to AFRL/SNAT, we are investigating FISST approaches for hedging against the uncertainties caused by poorly-understood sensor models (see section C.10). Under contract to MDA, we are implementing new computational approaches for multitarget detection and tracking (see sections C.6 and C.11). FISST robust track-fusion techniques are among those being investigated under contract to MRDEC (see section C.2).

B.8 UNPLANNED PROGRESS: RELATIONSHIP BETWEEN FISST AND POINT PROCESS THEORY

Point process theory is the stochastic theory of multi-object systems. It has two primary (and largely equivalent) mathematical formulations: in terms of stochastic geometry and random sets [4,96,115,131]; and in terms of random measure theory [13,39,105,121]. Finite-set statistics (FISST) is essentially a judicious and "engineering friendly" distillation of aspects of point process theory, expressed in the language of stochastic geometry. For example, the FISST concept of a "random finite set" is the same thing as a "simple point process." As already noted in the discussion in section B.6.2 above, all of the three following items are interchangeably known as a (simple) *multi-dimensional point process* [131, pp. 100-102], [4,13,115]: the random subset $\Xi_{k|k}$, the random "counting measure" $\Delta_{\Xi_{k|k}}(S)$, and its corresponding random density function $\delta_{\Xi_{k|k}}(\mathbf{x})$. Here,

$$\delta_X(\mathbf{x}) = \delta_{\mathbf{x}_1}(\mathbf{x}) + ... + \delta_{\mathbf{x}_n}(\mathbf{x})$$

$$\Delta_X(S) = \Delta_{\mathbf{x}_1}(S) + ... + \Delta_{\mathbf{x}_n}(S)$$

where $\delta_{\mathbf{w}}(\mathbf{x})$ denotes the Dirac delta function concentrated at $\mathbf{w}$ and where $\Delta_{\mathbf{w}}(\mathbf{x})$ denotes its corresponding Dirac measure: $\Delta_{\mathbf{w}}(S) = 1$ if $\mathbf{x} \in S$ and $\Delta_{\mathbf{w}}(S) = 0$ otherwise. Also, LMTS has known since 1995 that the FISST concept of a multitarget posterior, or multitarget likelihood function, is the same thing as the point process concept of a family of Janossy densities. Likewise, we have known since 1995 that the concept of a "probability hypothesis density" (PHD) is the same thing as the point process concept of a first-order moment measure [24, pp. 168-170], and we showed that the point process concept of a factorial-moment measure is the same thing as the FISST concept of a multitarget moment density function:

$$M_\Gamma(X) = \left[\frac{\delta \beta_\Gamma}{\delta X}(S) \right]_{S=S_{tot}}$$

where S_{tot} is the entire space of which S is a subset. Because of the work on multitarget moments (section B.6.2), which draws upon the point process concept of a moment measure, LMTS decided to extend our earlier work and clarify the relationship between FISST and the measure-theoretic version of point process theory. This work is described more fully in [59, pp.139-146].

In single-object statistics, the statistical behavior of a random number Y is often described by *generating functions* such as the characteristic function $\phi_Y(y) = E[e^{iyY}]$, the moment-generating function $M_Y(y) = E[e^{yY}]$, the factorial moment-generating function $G_Y(y) = E[y^Y]$, [10, p. 83], [13]. These functions are called "generating functions" because probability functions and various kinds of moments can be generated from their iterated derivatives, e.g. $(d^j M_Y / dy^j)(0) = E[Y^n]$. In like manner, the statistical behavior of the random finite set (simple point process) Γ is described by the *probability generating functional* (p.g.fl.) $G_\Gamma[h]$ [13]. Given a function h this is the expected value of the product of all $h(\mathbf{x})$ taken over all elements $\mathbf{x}$ of Γ:

$$G_\Gamma[h] = E\left[\prod_{\mathbf{x} \in \Gamma} h(\mathbf{x}) \right]$$

(This expectation will exist if, for example, $|h(\mathbf{x})|$ is bounded almost everywhere.) Probability generating functionals have the important property that $G_{\Gamma_1 \cup ... \cup \Gamma_n} = G_{\Gamma_1} \cdots G_{\Gamma_n}$ if $\Gamma_1,..., \Gamma_n$ are statistically independent random finite subsets.

One can differentiate p.g.fl.'s in essentially the same way that one differentiates ordinary functions. The Frechét functional derivative of a functional $F[h]$, with respect to an almost everywhere bounded function g, is the Dini differential quotient

$$\frac{\partial F}{\partial g}[h] = \lim_{\varepsilon \to 0} \frac{F[h + \varepsilon g] - F[h]}{\varepsilon}$$

if the limit exists and if it is linear and continuous in the argument g. Also, write

$$\frac{\partial^n F}{\partial g_1 \cdots \partial g_n}[h] = \frac{\partial}{\partial g_1} \cdots \frac{\partial}{\partial g_n} F[h]$$

The basic relationship between FISST and the random measure version of point process theory is as follows. First, given a subset S let 1_S be the *set-indicator* function defined by $1_S(\mathbf{x}) = 1$ if $\mathbf{x} \in S$ and $1_S(\mathbf{x}) = 0$ otherwise. Then the FISST belief-mass function $\beta_\Gamma(S)$ is the restriction of the corresponding probability generating functional $G_\Gamma[h]$ to the set-indicator functions:

$$\beta_\Gamma(S) = G_\Gamma[1_S]$$

Likewise, let $Y = \{\mathbf{y}_1, \ldots, \mathbf{y}_n\}$. Then the FISST set derivative of the belief-mass function $\beta_\Gamma(S)$ is the restriction of the iterated functional derivative of the corresponding probability generating functional $G_\Gamma[h]$ to the set-indicator functions:

$$\frac{\delta \beta_\Gamma}{\delta Y}(S) = \frac{\partial^n G_\Gamma}{\partial \delta_{\mathbf{y}_1} \cdots \partial \delta_{\mathbf{y}_n}}[1_S]$$

where $\delta_\mathbf{y}(\mathbf{x})$ denotes the Dirac delta function concentrated at the vector $\mathbf{y}$ and where the iterated derivative on the right is known in physics as a functional derivative [116, p. 173-174].

B.9 UNPLANNED PROGRESS: POINT TARGET-CLUSTERS AND CONTINUITY OF MULTITARGET DENSITY FUNCTIONS

In conventional single-sensor, single-target statistics, many techniques depend on the ability to apply Newtonian calculus techniques to the posterior density $f(\mathbf{x})$. For example, such techniques often assume that first and/or higher-order derivatives of $f(\mathbf{x})$ exist (see, for example, section B.10 below). Unfortunately, such techniques cannot be directly generalized to multitarget situations, because the multitarget posterior $f(X)$ is *inherently discontinuous with respect to changes in target number*. That is, the variable X experiences discontinuous jumps in its number of elements. During the Phase II contract, we initiated a study to determine ways to extend the variable X to more general state variables $\mathcal{X}$, and the multitarget posterior $f(X)$ to a more general posterior $f(\mathcal{X})$, in such a manner that $f(\mathcal{X})$ is at least continuous, and preferably differentiable, in the variable $\mathcal{X}$. Our preliminary results are summarized in [59, pp. 140, 158-159].

Briefly, we extend the concept of a point target with state vector $\mathbf{x}$ to that of a *point target-cluster* with state $(a, \mathbf{x})$. By this, we mean a cluster of unresolved targets, all of which are co-located at target-state $\mathbf{x}$, and the expected number of which is $a > 0$. If $a = 1$ then the point cluster $(1, \mathbf{x})$ models a single point target. A group of point clusters is just a finite set of the form $\mathcal{X} = \{(a_1, \mathbf{x}_1), \ldots, (a_n, \mathbf{x}_n)\}$. It can be shown that $\mathcal{X}$ is mathematically equivalent (in a point process sense) to the Dirac mixture density:

$$\{(a_1, \mathbf{x}_1), \ldots, (a_n, \mathbf{x}_n)\} \Leftrightarrow h = a_1 \delta_{\mathbf{x}_1} + \ldots + a_n \delta_{\mathbf{x}_n}$$

This identification allows us, in turn, to interpret *any* bounded nonnegative-valued function h as a *continuously infinite collection of point target-clusters*—meaning that a point track-cluster is located at each point $\mathbf{x}$ of state space and that the expected number-density of targets in this cluster is $h(\mathbf{x})$.

Given this, we can use the concept of a functional derivative (see section B.8) to differentiate functions of the form $f[h]$ and, in particular, functions of the form $f(\mathcal{X})$. The only question, then, is how to extend a multitarget density $f(X)$ defined on target state-sets X to a multitarget density $f(\mathcal{X})$ defined on target cluster-sets $\mathcal{X}$; or more generally, to a multitarget functional $f[h]$. We have shown how to define such

extensions by assuming that point clusters are independent of each other, and that each point-cluster has *Poisson statistics*. For example,

$$f(\{(a_1,\mathbf{x}_1),...,(a_n,\mathbf{x}_n)\}) = e^{-\alpha(a_1+...+a_n)} a_1 \cdots a_n \alpha^n f(\mathbf{x}_1) \cdots f(\mathbf{x}_n)$$

for some $\alpha > 0$.

B.10 UNPLANNED PROGRESS: MULTITARGET FUNCTIONALS AND "MULTITARGET KALMAN FILTER" AND "MULTITARGET EXTENDED KALMAN FILTER"

The point-cluster concepts summarized in section B.9 have been applied to an initial study of the possibility of constructing second-order approximate multitarget filters—that is, multitarget analogs of the extended Kalman filter (EKF). Some initial results were obtained and are reported in **[59, pp. 158-160]**.

B.10.1 Multitarget Covariance Densities. This analysis is partly based on the fact that *multitarget covariance functionals* can be defined for finite random sets. That is, let Ξ be a random finite set and let $\beta_\Xi(S) = \text{Pr}(\Xi \subseteq S)$ be its belief-mass function (section A.2.3). Then the *multitarget covariance density* of Ξ is the multitarget density defined by

$$c_\Xi(X) = \left[\frac{\delta \log \beta_\Xi}{\delta X}(S) \right]_{S=S_{tot}}$$

where S_{tot} is the entire space of which S is a subset. (In the conventional point process literature it is known as the family of *factorial cumulant densities* **[13]**.) This provides a potential basis for constructing multitarget analogs of the Kalman filter, by finding ways of propagating both the first order-moment (the PHD of section B.6.2) and the covariance moment. However, we did not find any evidence that this can be actually accomplished in a practical manner.

B.10.2 Multitarget Extended Kalman Filter. In the single-sensor, single-target case, the usual development of the EKF begins with nonlinear measurement and motion models **[12]**:

$$\mathbf{Z}_k = \mathbf{g}_k(\mathbf{x}_k) + \mathbf{V}_k, \quad \mathbf{X}_{k+1} = \mathbf{f}_k(\mathbf{x}_k) + \mathbf{W}_k$$

These models are then linearized:

$$\mathbf{g}_k(\mathbf{x}) \cong \mathbf{g}_k(\mathbf{x}_{k|k-1}) + \frac{\partial \mathbf{g}_k}{\partial \mathbf{x}}(\mathbf{x}_{k|k-1}) \cdot (\mathbf{x} - \mathbf{x}_{k|k-1})$$

$$\mathbf{f}_k(\mathbf{x}) \cong \mathbf{f}_k(\mathbf{x}_{k|k}) + \frac{\partial \mathbf{f}_k}{\partial \mathbf{x}}(\mathbf{x}_{k|k}) \cdot (\mathbf{x} - \mathbf{x}_{k|k})$$

where the partial derivatives indicate the Jacobian matrix of the indicated vector transformation. In what follows, for the sake of simplicity we ignore the motion model. The above approximation can be expressed equivalently in terms of the corresponding likelihood function as follows:

$$\log L_{z_k}(\mathbf{x}) \cong \log L_{z_k}(\mathbf{x}_{k|k-1}) + \frac{\partial \log L_{z_k}}{\partial(\mathbf{x} - \mathbf{x}_{k|k-1})}(\mathbf{x}_{k|k-1}) + \frac{1}{2}\frac{\partial^2 \log L_{z_k}}{\partial(\mathbf{x} - \mathbf{x}_{k|k-1})^2}(\mathbf{x}_{k|k-1})$$

In like manner, suppose that we have extended the multitarget likelihood function $L_Z(X)$ to a multitarget likelihood functional $L_Z[h]$ in the manner described in section B.9, which is to say that the function h represents a finite set of point-target clusters:

$$h = a_1 \delta_{\mathbf{x}_1} + ... + a_n \delta_{\mathbf{x}_n} \Leftrightarrow \{(a_1,\mathbf{x}_1),...,(a_n,\mathbf{x}_n)\}$$

Then we can define an analogous Taylor's series expansion

$$\log L_{Z_k}[h] \cong \log L_{Z_k}[h_{k|k-1}] + \frac{\partial \log L_{Z_k}}{\partial(h - h_{k|k-1})}[h_{k|k-1}] + \frac{1}{2}\frac{\partial^2 \log L_{Z_k}}{\partial(h - h_{k|k-1})^2}[h_{k|k-1}]$$

In principle, a multitarget EKF can be built around this expansion, together with the analogous expansion for the multitarget Markov density. However, the resulting quantities will be functional expressions involving integral transforms of the functions h. It is unclear at this time whether or not the resulting "multitarget EKF" can be rendered computationally viable.

SECTION C: TECHNOLOGY TRANSITION

Though the peer reviews for our second contract were very positive, the reviewers also strongly recommended that FISST techniques be put to the test by applying them to real-world problems. Since a basic research contract budget stretches only so far, and since focusing on a single "pet rock technology" risks squandering scarce resources on a solution that nobody actually wants (despite all expectations to the contrary), LMTS addressed this issue as follows. We leveraged our Phase II USARO contract as basic-research "intellectual venture capital," using it to develop a range of innovative FISST-based techniques directed at a range of applications. This "omnidirectional" technology-leveraging strategy would, we believed, increase the likelihood that at least some techniques would attract the funding necessary to support application to real-world problems. The results of this strategy exceeded all expectations. The following FISST-derived DoD research contracts were won during the term of the Phase II contract. (Each contract is briefly described in turn in the indicated subsections below):

1. MRDEC: Sensor Data Fusion for Target Identification (beginning; section C.2)
   - track-level fusion of ID's produced by multiple HRRR classifier algorithms
2. AFRL/IFEA: Unified Metrology for Data Fusion (completed; section C.3)
   - scientific performance evaluation of Level 1 fusion algorithms using information theory MoEs
3. AFRL/IFEA: Measures of Effectiveness for Abstract Data Fusion (in progress; section C.8)
   - scientific performance evaluation of Levels 2, 3, 4 fusion algorithms using information theory MoEs
4. AFOSR: Unified Collection and Control for UCAV Swarms (beginning; section C.4)
   - fundamental concepts in sensor and platform control, and data fusion, for UCAVs
5. MDA: Project Hercules (beginning; section C.6)
   - first-principles approach for cluster target tracking
6. MDA: Unified Bayesian Cluster Tracking and Discrimination (beginning; section C.11)
   - first-principles approach for joint cluster tracking and discrimination
7. AFRL/SNAT: Space-Based Targeting Technologies (in progress; section C.7)
   - joint tracking, pose estimation, and identification via fusion of track & HRR radar data
8. AFRL/SNAT: Unified Evidence Accrual for Data Fusion (in progress; section C.10)
   - robust SAR ATR of stationary ground targets under Extended Operating Conditions

In addition, USARO-funded work has led to target identification techniques that have been used in the following programs:

9. SPAWAR Systems Center: Deployable Autonomous Distributed System (completed; section C.5)
10. LMTS internal research & development: ASW/ASUW Data Fusion Workstation (continuing);
11. LMTS internal research & development: C⁴I INTELL Robust Data Fusion & Target ID (completed);

Most of these contracts are just beginning, and so assessment of the utility of FISST techniques is premature at this time. However, the longest-running effort (contracts 2 and 3) has successfully demonstrated the potential utility of FISST information theory MoEs to multisensor-multitarget performance evaluation applications.

C.1 Project Title: Information-Theoretic Information Fusion. *Funding Agency:* USARO/ Dr. William Sander (919-549-4241). *Contracts:* DAAH04-94-C-0011, DAAG55-98-C-0039. *Contract type:* BAA. *Performance Period:* 1994 - 2001. *Description:* Under two consecutive three-year basic research contracts, LMTS is developing a unified, theoretically rigorous approach to data fusion based on information theory and finite-set statistics (FISST). *LMTS Principal Investigator:* Ronald Mahler.

C.2 Project Title: Sensor Data Fusion for Target Identification. *Funding Agency*: U.S. Army MRDEC/ Dr. Scott Holder (256-842-8997). *Contracts*: DAAH01-01-C-R110, TBD. *Contract type*: SBIR Phase I, Phase II (prime contractor, SSCI). *Performance Period*: 2001 - 2004. *Description*: LMTS, together with its small business partner SSCI, is developing methods for fusing the outputs of multiple target identification algorithms whose inputs are High Range Resolution Radar (HRRR) signatures. FISST-based identification fusion techniques are among those being investigated. *LMTS Principal Investigator:* Ronald Mahler. *References*: N/A; this is a new project.

C.3 Project Title: Unified Metrology for Data Fusion. *Funding Agency*: AFRL/IFEA/ Ms. Barbara Lajza-Rooks (315-330-3055). *Contract*: F30602-98-C-0270. *Contract Type*: BAA. *Performance Period*: 1998-2000. *Description*: LMTS developed a FISST-based, unified, theoretically defensible approach to metrology for multisource-multitarget Level 1 data fusion algorithms for application in performance evaluation, fusion management, and adaptive data fusion. *LMTS Principal Investigator:* Ronald Mahler. *References*: **[33,140,141]**.

C.4 Project Title: Unified Collection and Control for UCAV Swarms. *Funding Agency: AFOSR / Dr. Jon Sjogren (703-696-6564).* *Contract*: F49620-01-C-0031. *Contract Type*: BAA. *Performance Period*: 2001-2004. *Description:* LMTS and its subcontractor Scientific Systems Co. Inc. will be conducting basic research to develop fundamental approaches data fusion, sensor management, and platform management of swarms of Unattended Combat Aerial Vehicles (UCAVs). FISST approaches to data fusion and assets management will be a basic part of this program. *References:* N/A; this project is new.

C.5 Project Title: Deployable Autonomous Distributed System (DADS). *Funding Agency*: SPAWAR Systems Center/ Ms. Joan Kaina (619-553-2347). *Contracts*: N00039-95-C-0080, N00024-96-G-5207. *Performance Period*: 1994 - 2000. *Description*: LMTS and its subcontractor Summit Research Corp. developed an automatic algorithm to identify surface combatants and brown- and blue-water submarines directly from OTH-Gold features extracted from passive-acoustic, ELINT, magnetic field, and electric field data. The expert-systems technique used in this classifier makes use of certain aspects of the finite-set statistics (FISST) approach described in this proposal. *LMTS Principal Investigator:* Ronald Mahler. *References*: **[1,28,29]**

C.6 Project Title: Project Hercules. *Funding Agency*: BMDO/HQ / Lt. Col. James Myers (703-601-4219). *Contracts*: N00024-96-G-5207, WQ88, Lockheed Martin IWTA. *Contract Type*: BOA, TOA. *Performance Period*: 2000-ongoing. *Description*: Dr. Ronald Mahler is part of a team of nationally-known experts supporting the Project Hercules Advanced Technology Panel team and the Data Fusion Panel team in the development of new advanced techniques in detection, tracking, and discrimination for ballistic missile defense. LMTS is to develop new approaches for the tracking of RVs in clouds of countermeasures, based on FISST first-order multitarget moment statistic approximations. *LMTS Principal Investigator:* Ronald Mahler. *References*: N/A; this project is new.

C.7 Project Title: Space-Based Targeting Technologies. *Funding Agency*: AFRL/SNAT, Dayton OH/Mr. Michael Noviskey (937-255-1115 x3321). *Contracts:* F33615-99-C-1454, F33615-00-C-1616. *Contract Type*: SBIR Phase I, Phase II (prime contractor, SSCI). *Performance Period*: 1999 – 2002. *Description*: LMTS and its small business partner SSCI are developing an algorithm to jointly track and identify air targets by fusing High Range Resolution Radar (HRRR) data and tracking-radar data. This algorithm consists of a nonlinear filter (track data) coupled with HRRR classifier algorithms developed by AFRL/SNAT, SSCI, and LMTS. *LMTS Principal Investigator:* Ronald Mahler. *References*: **[93,142]**

C.8 Project Title: Measures of Effectiveness for Abstract Data Fusion. *Funding Agency*: AFRL/IFEA, Rome NY/Mr. Mark Alford (315-330-3573). *Contracts*: F30602-99-C-0124, F30602-00-C-

0085. *Contract Type*: SBIR Phase I, Phase II (prime contractor, SSCI). *Performance Period*: 1999 - 2002. *Description*: This is a follow-on program to the project described in section C.3. LMTS and its small business partner SSCI are developing a FISST-based scientific basis for performance evaluation for multisource-multitarget data fusion algorithms at Level 2 fusion (threat assessment), Level 3 fusion (threat assessment), and Level 4 fusion (sensor/asset management). *LMTS Principal Investigator:* Ronald Mahler. *References*: **[15,17]**.

C.9   Project Title: Tracking in High-Scintillation Environments. *Funding Agency*: AFRL/DEBA, Albuquerque NM/Dr. Donald Washburn (505-846-1597). *Contracts*: F29601-00-C-0091. *Contract Type*: SBIR Phase I, Phase II (prime contractor, SSCI). *Performance Period*: 1999-2002. *Description*: LMTS and its small business partner SSCI are developing approaches and algorithms for tracking missile and ground targets using Air Borne Laser (ABL) ladar data. *LMTS Principal Investigator:* Ronald Mahler. *References*: **[114]**

C.10   Project Title: Unified Evidence Accrual for Data Fusion. *Funding Agency:* AFRL/SNAT / Stanton Musick (937-255-1491 x4292). *Contract*: F33615-98-C-1292, F33615-99-C-1430. *Contract Type:* SBIR Phase I, Phase II (prime contractor, SSCI). *Performance Period*: 1998-2002. *Description:* LMTS and its small business partner SSCI are conducting applied research aimed at developing a theoretically defensible, robust approach to Automatic Target Recognition (ATR) for Synthetic Aperture Radar (SAR) against stationary ground targets under Extended Operating Conditions (EOC). FISST-based techniques are among those being investigated. *LMTS Principal Investigator:* Ronald Mahler. *References*: **[36,94]**

C.11   Project Title: Unified Bayesian Cluster Target Tracking and Discrimination. *Funding Agency*: BMDO / Mr. Alexander Gilmore (256-955-1568). *Contract*: DASG60-01-P-0032, TBD. *Contract Type*: SBIR Phase I,II (prime contractor, SSCI). *Performance Period*: 2001-2002. *Description*: LMTS and its small business partner are to develop a preliminary FISST-based approach for achieving joint tracking and discrimination of RVs in clouds of countermeasures. *LMTS Principal Investigator:* Ronald Mahler. *References*: N/A; this project is new.

C.12   Project Title: New Non-Cooperative Target Recognition Techniques. *Funding Agency*: NAVAIR / Mr. George Linde (202-767-2643). *Contract*: N00024-01-C-4071. *Contract Type*: SBIR Phase I (prime contractor, SSCI). *Performance Period*: 2001-2001. *Description*: LMTS and its small business partner are to develop a simulation environment and target classification algorithm for automatic target recognition using High Range Resolution (HRRR) signatures. *LMTS Principal Investigator:* Ronald Mahler. *References*: N/A; this project is new.

C.13   Project Title: Adaptive Data Fusion Technologies. *Funding Agency*: AFRL/IFEA, Rome NY/Ms. Barbara Lajza-Rooks (315-330-3055). *Contract*: F33615-98-C-1292. *Contract Type*: Phase I SBIR (prime contractor, SSCI). *Performance Period*: 1998-1998. *Description*: LMTS and its small business partner SSCI developed a FISST-based, self-reconfiguring adaptive, robust target identification algorithm based on multisource datalink attributes in USMTF message format. (This project did not go onto a Phase II). *LMTS Principal Investigator:* Ronald Mahler. *References*: **[14]**

C.14   Project Title: Innovative Information Technologies. *Funding Agency*: AFRL/IFEA, Rome NY/Ms. Barbara Lajza-Rooks (315-330-3055). *Contract*: F30602-00-C-0107. *Contract Type*: Phase I SBIR (prime contractor, SSCI). *Performance Period*: 2000-2000. *Description*: LMTS and its small business partner SSCI developed a preliminary, FISST-based implementation of a multitarget tracking and identification filter based on the concept of a multitarget first-order moment statistic. (This project did not go onto a Phase II). *LMTS Principal Investigator:* Ronald Mahler. *References*: **[16]**

C.15 Project Title: C4I Data Fusion Technologies. *Funding Agency*: LMTS, Eagan MN. *Contract*: IR&D. *Performance Period*: 1998-2000. *Description*: This was an LMTS, internally-funded follow-on to the project described in section C.13. LMTS and its small business partner Summit Research Corp. developed a FISST-based, self-reconfiguring adaptive, robust target identification algorithm based on multisource datalink attributes in USMTF message format. *LMTS Principal Investigator:* Ronald Mahler. *References*: **[122]**

SECTION D: PROJECT-GENERATED PUBLICATIONS

See section G below for specific reference information regarding the following publications.

Monograph
R. Mahler (2000) *An Introduction to Multisource-Multitarget Statistics and Its Applications*

Chapters in Books
R. Mahler (2001) "Random Set Theory for Target Tracking and Identification"

Papers Submitted to Journals
R. Mahler (2000) "Approximate Multisensor-Multitarget Joint Detection, Tracking, and Identification Using a First-Order Multitarget Moment Statistic," currently in review

Conference Papers
1. R. Mahler (2001) "Detecting, tracking, and classifying group targets: A unified approach,"
2. R. Mahler (2001) "Engineering Statistics for Multi-Object Tracking"
3. R. Mahler (2001) "Multitarget filtering using a multitarget first-order moment statistic"
4. R. Mahler (2001) "Multitarget moments and their application to multitarget tracking"
5. R. Mahler (2000) "Optimal/robust distributed data fusion: a unified approach"
6. R. Mahler (2000) "The search for tractable Bayesian multitarget filters"
7. R. Mahler (2000) "A theoretical foundation for the Stein-Winter 'Probability Hypothesis Density (PHD)' multitarget tracking approach"
8. R. Mahler (1999) "Multitarget Detection and Acquisition: A Unified Approach"
9. R. Mahler (1999) "Multitarget Markov motion models"
10. R. Mahler (1999) "Why Multi-Source, Multi-Target Data Fusion is Tricky"
11. R. Mahler and M. O'Hely (1999) "Multitarget detection and acquisition: A unified approach"
12. R. Mahler (1998) "Global posterior densities for sensor management"
13. R. Mahler (1998) "Information for fusion management and performance estimation"
14. R. Mahler (1998) "Multisource, multitarget filtering: A unified approach"

SECTION E: PARTICIPATING SCIENTIFIC PERSONNEL

Ronald P.S. Mahler, Ph.D.

SECTION F: PROJECT-GENERATED INVENTIONS

None.

SECTION G: BIBLIOGRAPHY

1. R. Allen, R. Myre, J. Warner, J. Hatlestad, R. Mahler, P. Ohmann, T.C. Poling, E. Taipale, and W. Richter (1998) "Passive-Acoustic Classification System (PACS) for ASW, *Proc. 1998 IRIS Nat'l Symp. on Sensor and Data Fusion*, Vol. I (Unclassified), Mar. 31 - Apr. 2, Marietta GA, pp. 179-192

2. D.L. Alspach (1975) ``A Gaussian Sum Approach to the Multi-Target Identification-Tracking Problem,`` *Automatica*, Vol. 11, pp. 285-296

3. D.J. Ballantyne, H.Y. Chan, and M.A. Kouritzin (2001) "A branching particle-based nonlinear filter for multi-target tracking," *Proc. 2001Int'l Conf. on Information Fusion*, Aug. 7-10 2001, Montreal, to appear

4. M. Bardin (1984) ``Multidimensional Point Processes and Random Closed Sets," *J. Applied Prob.*, pp. 173-178

5. C.A. Barlow, L.D. Stone, and M.V. Finn (1996) "Unified Data Fusion," *Proc. 9'th Nat'l Symp. on Sensor Fusion*, Vol. I (Unclassified), Naval Postgraduate School, Monterey CA, Mar. 11-13 1996, pp. 321-330

6. Y. Bar-Shalom and X.-R. Li (1993) *Estimation and Tracking: Principles, Techniques, and Software*, Artech House

7. V.E. Beneš (1981) "Exact finite-dimensional filters for certain diffusions with nonlinear drift," *Stochastics*, Vol. 5, pp. 65-92

8. R.E. Bethel and G.J. Paras (1994) "A PDF multitarget-tracker," *IEEE Trans AES*, Vol. 30 No. 2, pp. 386-403

9. D.E. Brown and H.L. Liu (2001) "A point process transition density model for threat assessment," *Proc. 2001 MSS Nat'l Symp. on Sensor and Data Fusion*, Vol. I (Unclassified), June 25-28 2001, San Diego, to appear

10. G. Casella and R.L. Berger (1990) *Statistical Inference*, Wadsworth & Brooks

11. S. Challa, Y. Bar-Shalom, and V. Krishnamurthy (2000) "Nonlinear filtering via generalized Edgeworth series and Gauss-Hermite Quadrature," *IEEE Trans. Signal Processing*, vol. 48 no. 6, pp. 1816-1820

12. C.K. Chui and G. Chen (1999) *Kalman Filtering With Real-Time Applications*, Third Edition, Springer-Verlag

13. D.J. Daley and D. Vere-Jones (1988) *An Introduction to the Theory of Point Processes*, Springer-Verlag

14. A.El-Fellah, R. Mahler, B. Ravichandran, and R. Mehra (1999) "Adaptive Data Fusion Using Finite-Set Statistics," in I. Kadar (ed.), *Signal Processing, Sensor Fusion, and Target Recognition VIII.*, SPIE Vol. 3720, pp. 80-91, ISBN 0-8194-3194-X

15. A. El-Fallah, R. Mahler, T. Zajic, E. Sorensen, M. Alford, and R. Mehra (2000) "Scientific performance evaluation for sensor management," in I. Kadar (ed.), *Signal Processing, Sensor Fusion, and Target Recognition IX*, SPIE Vol. 4052, pp. 183-194, ISBN 0-8194-3678-X

16. A. El-Fallah, T. Zajic, R. Mahler, B. Ravichandran, and R. Mehra (2001) "Multitarget nonlinear filtering based on spectral compression and probability hypothesis density," in I. Kadar (ed.), *Signal Processing, Sensor Fusion, and Target Recognition X*, SPIE Vol. 4380, pp. 207-216, ISBN 0-8194-4075-2

17. A. El-Fallah, J. Hoffman, T. Zajic, R. Mahler, and R. Mehra (2001) "Scientific performance evaluationfor distributed sensor management and adaptive data fusion," in I. Kadar (ed.), *Signal Processing, Sensor Fusion, and Target Recognition X*, SPIE Vol. 4380, pp. 328-338, ISBN 0-8194-4075-2

18. D. Fixsen and R. Mahler (1997) "The modified Dempster-Shafer approach to classification," *IEEE Trans. SMC-Part A*, vol. 27 no. 1, pp. 96-104

19. C.A.J. Fletcher (1988) *Computational Techniques for Fluid Dynamics: Fundamental and General Techniques, Vol. 1*, Springer-Verlag

20. I.M. Gel'fand and A.M. Yaglom (1960), "Integration in Function Spaces and its Applications in Quantum Physics," *J. of Math. Physics*, Vol. 1 No. 1, pp. 48-560

21. I.R. Goodman (1982) ``Fuzzy sets as equivalence classes of random sets,`` in R. Yager (ed.), *Fuzzy Sets and Possibility Theory*, Permagon, pp. 327-343

22. I.R. Goodman (1994) "Toward a Comprehensive Theory of Linguistic and Probabilistic Evidence: Two New Approaches to Conditional Event Algebra," *IEEE Trans. on Sys., Man, and Cybern.*, Vol. 24, pp. 1865-1698

23. I.R. Goodman, R.P.S. Mahler, and H.T. Nguyen (1999) "What is Conditional Event Algebra and Why Should You Care?", in I. Kadar (ed.) *Signal Processing, Sensor Fusion, and Target Recognition VIII*, SPIE Vol. 3720, pp. 2-13

24. I.R. Goodman, R.P.S. Mahler, and H.T. Nguyen (1997) *Mathematics of Data Fusion*, Kluwer Academic Publishers, ISBN 0-7923-4674-2

25. I.R. Goodman and H.T. Nguyen (1985) *Uncertainty Models for Knowledge Based Systems*, North-Holland

26. J. Goutsias, R.P.S. Mahler, and H.T. Nguyen, eds. (1997) *Random Sets: Theory and Applications*, Springer-Verlag, ISBN 0-3879-8345-7

27. M. Grabisch, H.T. Nguyen, and E.A. Walker (1995), *Fundamentals of Uncertainty Calculi With Applications to Fuzzy Inference*, Kluwer Academic Publishers

28. M. Hatch, J.L. Kaina, R.P. Mahler, and R.S. Myre (1998) "Data Fusion Methodologies to Support Theater Level and Deployable Surveillance Systems," *Proc. 1998 Asilomar Conf. on Signals, Systems, and Computers*, Naval Postgraduate School, Monterey CA, Nov. 1-3 1998, pp. 563-567

29. M. Hatch, J.L. Kaina, M. Owen, R.P. Mahler, R.S. Myre, and S.J. Benkoski (1998) "Data Fusion Methodologies in the Deployable Autonomous Distributed Systems (DADS) Project, *Proc. Int'l Conf. on Information Fusion*, Las Vegas, to appear

30. K. Hestir, H.T. Nguyen, and G.S. Rogers (1991) ``A random set formalism for evidential reasoning,`` in I.R. Goodman, M.M. Gupta, H.T. Nguyen and G.S. Rogers (eds.), *Conditional Logic in Expert Systems*, North-Holland, pp. 309-344

31. T.L.Hill (1956) *Statistical Mechanics: Principles and Practical Applications*, Dover Publications 1987

32. Y.C. Ho and R.C.K. Lee (1964) ``A Bayesian Approach to Problems in Stochastic Estimation and Control," *IEEE Trans. Automatic Contr.*, Vol. 9, pp. 333-339

33. J. Hoffman, R. Mahler, and T. Zajic (2001) "User-defined information and scientific performance evaluation," in I. Kadar (ed.), *Signal Processing, Sensor Fusion, and Target Recognition X*, SPIE Vol. 4380, pp. 300-311, ISBN 0-8194-4075-2

34. A. Höhle (1981) "A mathematical theory of uncertainty: fuzzy experiments and their realizations," in R.R. Yager (ed.), *Recent Developments in Fuzzy Set and Possibility Theory*, Permagon Press, pp. 344-355

35. P.J. Huber (1981) *Robust Statistics*, John Wiley & Sons

36. M. Huff, S.-H. Yu, R. Mahler, B. Ravichandran, R.K. Mehra, and S. Musick, "Unified evidence accrual for SAR: Recent results," in I. Kadar (ed.), *Signal Processing, Sensor Fusion, and Target Recognition IX*, SPIE Vol. 4052, pp. 149-159, ISBN 0-8194-3678-X

37. R.A. Iltis (1999) "State Estimation Using an Approximate Reduced Statistics Algorithm," *IEEE Trans. Aerospace and Electr. Sys.*, vol. 35 no. 4, pp. 1161-1172

38. A.H. Jazwinski (1970) *Stochastic Processes and Filtering Theory*, Academic Press

39. A.F. Karr (1991) *Point Processes and Their Statistical Inference*, Second Edition, Marcel Dekker

40. K. Kastella (1997) "Joint multitarget probabilities for detection and tracking," in M.K. Masten and L.A. Stockum (eds.), *Acquisition, Tracking, and Pointing XI*, SPIE Vol. 3086, pp. 122-128, ISBN 0-8194-2501-X

41. K. Kastella (2000) "Finite difference methods for nonlinear filtering and automated target recognition," in Y. Bar-Shalom and W.D. Blair (eds.), *Multitarget-Multisensor Tracking: Applications and Advances, Vol. III*, Artech House, 2000, pp. 233-258

42. K. Kastella (2000) "A microdensity approach to multitarget tracking," *Proc. 2000 Int'l Conf. on Information Fusion*, Paris

43. K. Kastella (1996) "Discrimination gain for sensor management in multitarget detection and tracking," *Proc. 1996 IMACS Multiconf. on Comp. and Eng. Appl. (CESA'96): Symp. on Contr., Opt., and Supervision*, Lille France, July 9-12 1996, pp. 167-172

44. K. Kastella, M.A. Kouritzin, and A. Zatezalo (1996) "A nonlinear filter for altitude tracking," *Proc. 41st Annual Int'l Program and Exhib. of the Air Traffic Contr. Assoc.*, Nashville TN, Oct. 13-17 1996, pp. 152-156

45. P.A. Kelly, H. Derin, and P. Vaidya (2000) "Qualitative optimization of image processing systems using random set modeling," in I. Kadar (ed.), *Signal Processing, Sensor Fusion, and Target Recognition IX*, SPIE Vol. 4052, pp. 139-148, ISBN 0-8194-3678-X

46. M.A. Kouritzin (2001) "Particle Approximations," presentation at the AFRL/AFOSR Workshop on Nonlinear Filtering Methods for Tracking, Dayton OH, Feb. 21-22

47. M.A. Kouritzin (1998) "On exact filters for continuous signals with discrete observations," *IEEE Trans. Auto. Control*, vol. 43, pp. 709-715

48. M.A. Kouritzin (1996) "On exact filters for continuous signals with discrete observations", Technical Report No. 1409, University of Minnesota Institute for Mathematics and Its Applications, May 1996

49. R. Kruse, E. Schwencke, and J. Heinsohn (1991*) Uncertainty and Vagueness in Knowledge-Based Systems*, Springer-Verlag

50. R. Kulhavý (1990) "Recursive Nonlinear Estimation: A Geometric Approach," *Automatica*, vol. 26 no. 3, pp. 545-555

51. A.D. Lanterman, M.I. Miller, D.L. Snyder, and W.J. Miceli (1994) "Jump-diffusion processes for the automated understanding of FLIR scenes," in F.A. Sadjadi (ed.) *Automatic Target Recognition IV*, SPIE Vol. 2234, pp. 416-427, ISBN 0-8194-1538-3

52. P.M. Lee (1997) *Bayesian Statistics: An Introduction*, Second Edition, Arnold Publishers/John Wiley

53. X.-R. Li (1999) "Tracking in the Presence of Range Deception ECM and Clutter by Decomposition and Fusion," in O.E. Drummond (ed.) *Tracking and Signal Processing of Small Targets 2000*, SPIE Vol. 3809, pp. 198-210, ISBN 0-8194-3295-4

54. Y. Li (1994) "Probabilistic Interpretations of Fuzzy Sets and Systems," Doctoral Dissertation, Dept. of Elec. Eng. and Comp. Sci., M.I.T., July 1994

55. R. Lototsky, R. Mikulevicius, and B.L. Rozovskii (1997) "Nonlinear filtering revisited: A spectral approach," *SIAM J. Contr. Optim.*, Vol, 35, pp. 435-461

56. R. Mahler (2001) "Detecting, tracking, and classifying group targets: A unified approach," in I. Kadar (ed.), *Signal Processing, Sensor Fusion, and Target Recognition X*, SPIE Vol. 4380, pp. 217-228, ISBN 0-8194-4075-2

57. R. Mahler (2001) "Engineering Statistics for Multi-Object Tracking," *Proc. 2001 IEEE Workshop on Multi-Object Tracking*, July 8, 2001, Vancouver, pp. 53-60, ISBN 0-7695-1171-6

58. R. Mahler (2001) "Multitarget filtering using a multitarget first-order moment statistic," in I. Kadar (ed.), *Signal Processing, Sensor Fusion, and Target Recognition X*, SPIE Vol. 4380, pp. 184-195, ISBN 0-8194-4075-2

59. R. Mahler (2001) "Multitarget moments and their application to multitarget tracking," *Proc. Workshop on Estimation, Tracking, and Fusion: A Tribute to Yaakov Bar-Shalom*, May 17, 2001, Naval Postgraduate School, Monterey CA, pp. 134-166, ISBN 0-9648-3124-4

60. R. Mahler (2001) "Random Set Theory for Target Tracking and Identification," in D.L. Hall and J. Llinas (eds.), *Handbook of Multisensor Data Fusion*, CRC Press, Boca Raton FL, pp. 14-1 to 14-133, ISBN 0-8493-2379-7

61. R. Mahler (2000) "Approximate Multisensor-Multitarget Joint Detection, Tracking, and Identification Using a First-Order Multitarget Moment Statistic," submitted to *IEEE Trans. AES*

62. R. Mahler (2000) *An Introduction to Multisource-Multitarget Statistics and Its Applications*, Lockheed Martin Technical Monograph, Mar. 15 2000, 114 pages

63. R. Mahler (2000) "Optimal/robust distributed data fusion: a unified approach," in I. Kadar (ed.), *Signal Processing, Sensor Fusion, and Target Recognition IX*, SPIE Vol. 4052, pp. 128-138, ISBN 0-8194-3678-X

64. R. Mahler (2000) "The search for tractable Bayesian multitarget filters," in O.E. Drummond (ed.) *Signal and Data Processing of Small Targets.*, SPIE Vol. 4048, pp. 310-320, ISBN 0-8194-3674-7

65. R. Mahler (2000) "A theoretical foundation for the Stein-Winter 'Probability Hypothesis Density (PHD)' multitarget tracking approach," *Proc. 2000 MSS Nat'l Symp. On Sensor and Data Fusion*, Vol. I (Unclassified), Kelly AFB, San Antonio TX, June 20-22 2000, DTIC/IRIA Oct. 2000, 440000-185-X(I), pp. 99-118

66. R. Mahler (1999) "Multitarget Detection and Acquisition: A Unified Approach," in O.E. Drummond (ed.) *Signal and Data Processing of Small Targets 1999.*, SPIE Vol. 3809, pp. 218-229, ISBN 0-8194-3295-4

67. **67a.** R. Mahler (1999) "Multitarget Markov motion models," in I. Kadar (ed.) *Signal Processing, Sensor Fusion, and Target Recognition VIII.*, SPIE Vol. 3720, pp. 47-58, ISBN 0-8194-3194-X; **67b.** R. Mahler (1999) "Why Multi-Source, Multi-Target Data Fusion is Tricky," *Proc. 1999 IRIS Na'l Symp. on Sensor and Data Fusion*, Vol. I (Unclassified), Johns Hopkins Applied Physics Laboratories, Laurel MD, pp. 135-153

68. R. Mahler (1998) "Global posterior densities for sensor management," in M.K. Kasten and L.A. Stockum (eds.), *Acquisition, Tracking, and Pointing XII.*, SPIE Vol. 3365, pp. 252-263, ISBN 0-8194-2814-0

69. R. Mahler (1998) "Information for fusion management and performance estimation," in I. Kadar (ed.) *Signal Processing, Sensor Fusion, and Target Recognition VII.*, SPIE Vol. 3374, pp. 64-75, ISBN 0-8194-2823

70. R. Mahler (1998) "Multisource, multitarget filtering: A unified approach," in O.E. Drummond (ed.) *Signal and Data Processing of Small Targets 1998*, SPIE Vol. 3373, pp. 296-307, ISBN 0-8194-2822-1

71. R. Mahler (1998) "What is the 'Multitarget Average Value' of a Multitarget Scenario," presentation at 1998 GTRI/ONR Workshop on Tracking and Filtering, Georgia Tech Research Institute, Marietta GA, June 1 1998

72. R. Mahler (1997) "Decisions and Data Fusion," *Proc. 1997 Nat'l Symp. on Sensor and Data Fusion*, April 14-17 1997, M.I.T. Lincoln Laboratories, pp. 71-87

73. R. Mahler (1997) "Measurement models for ambiguous evidence using conditional random sets," in I. Kadar (ed.) *Signal Processing, Sensor Fusion, and Target Recognition VI.*, SPIE Vol. 3068, pp. 40-51, ISBN 0-8194-2483-8

74. R. Mahler (1997) "Random sets in information fusion: An Overview," in J. Goutsias, R.P.S. Mahler, and H.T. Nguyen (eds.), *Random Sets: Theory and Applications*, Springer-Verlag, pp. 129-164, ISBN0-3879-8345-7

75. R. Mahler (1997) "A Theoretical Unification of Knowledge-Based Systems With Multisensor, Multitarget Estimation," in P. Wang (ed.) *Advances in Machine Intelligence and Soft Computing, Vol. IV*, Duke University Dept. of Elec. Eng.

76. R. Mahler (1996) "Combining Ambiguous Evidence With Respect to Ambiguous *a priori* Knowledge, I: Boolean Logic," *IEEE Trans. SMC-art A*, vol. 26, pp. 27-41

77. R. Mahler (1996) "Global Optimal Sensor Allocation," *Proc. Ninth Nat'l Symp. on Sensor Fusion*, Vol. I (Unclassified), Mar. 12-14 1996, Naval Postgraduate School, Monterey CA, pp. 347-366

78. R. Mahler (1996) "Representing Rules as Random Sets: Statistical Correlations Between Rules," *Information Sciences*, vol. 88, pp. 47-68

79. R. Mahler (1996) "Representing Rules as Random Sets, II: Iterated Rules," *Inn'l Jour. Intell. Sys.*, vol. 11, pp. 583-610

80. R. Mahler (1996) "Unified data fusion: fuzzy logic, evidence, and rules," in I. Kadar (ed.) *Signal Processing, Sensor Fusion, and Target Recognition V*, SPIE Vol. 2755, pp. 226-237, ISBN 0-8194-2136-7

81. R. Mahler (1995) "Combining Ambiguous Evidence With Respect to Ambiguous *a priori* Knowledge, II: Fuzzy Logic," *Fuzzy Sets and Systems*, vol. 75, pp. 319-354

82. R. Mahler (1995) "Finite-set statistics with application to data fusion," in A. Friedman, ed., *Mathematics in Industrial Problems, Part 7*, Springer-Verlag, pp. 198-206, ISBN 0-3879-4444-3

83. R. Mahler (1995) "Information Theory and Data Fusion," *Proc. Eighth Nat'l Symp. on Sensor Fusion*, Vol. I (Unclassified), Dallas TX, March 15-17 1995, ERIM, Ann Arbor, pp. 279-292

84. R. Mahler (1995) Nonadditive probability, finite-set statistics, and information fusion," *Proc. 34th IEEE Conf. on Decision and Control*, New Orleans, Dec. 13-15 1995, pp. 1947-1952

85. R. Mahler (1995) "Unified nonparametric data fusion," in I. Kadar (ed.), *Signal Processing, Sensor Fusion, and Target Recognition IV*, SPIE Vol. 2484, pp. 66-74, ISBN 0-8194-1837-4

86. R. Mahler (1994) "Classification when *a priori* evidence is ambiguous," in F.A. Sadjadi (ed.) *Automatic Object Recognition IV*, Vol. 2234, pp. 296-304, ISBN 0-8194-1538-3

87. R. Mahler (1994) "Global Integrated Data Fusion," *Proc. 7th Nat'l Symp. on Sensor Fusion*, Vol. I (Unclassified), Sandia National Laboratories, Albuquerque NM, March 16-18 1994, ERIM, Ann Arbor, pp. 187-199

88. R. Mahler (1994), "The random set approach to data fusion," in F.A. Sadjadi (ed.) *Automatic Object Recognition IV.*, SPIE Vol. 2234, pp. 287-295, ISBN 0-8194-1538-3

89. R. Mahler (1994) "Systematic data fusion using the theory of conditional random sets," in A. Friedman, ed., *Mathematics in Industrial Problems, Part 6*, pp. 156-165, ISBN 0-3879-4157-6

90. R. Mahler (1994/1996) "A unified foundation for data fusion," in F.A. Sadjadi (ed.), *Selected Papers on Sensor and Data Fusion*, SPIE Vol. MS-124, 1996, pp. 325-345; reprinted from *Seventh Joint Service Data Fusion Symposium*, Laurel MD, 1994, pp. 154-174, ISBN 0-8194-2265-7

91. R. Mahler and M. O'Hely (1999) "Multitarget detection and acquisition: A unified approach," in O.E. Drummond (ed.), *Signal and Data Processing of Small Targets 1999*, SPIE Vol. 3809, pp. 218-229, ISBN 0-8194-3295-4

92. R. Mahler, P. Leavitt, J. Warner, and R. Myre (1999) "Nonlinear filtering with really bad data," in I. Kadar (ed.), *Signal Processing, Sensor Fusion, and Target Recognition VIII.*, Vol. 3720, pp. 59-70, ISBN 0-8194-3194-X

93. R. Mahler, C. Rago, T. Zajic, S. Musick, and R.K. Mehra (2000) "Joint tracking, pose estimation and identification using HRRR data," in I. Kadar (ed.), *Signal Processing, Sensor Fusion, and Target Recognition IX*, SPIE Vol. 4052, 195-206, ISBN 0-8194-3678-X

94. R. Mahler, S.-H. Yu, R.K. Mehra, B. Ravichandran, and S. Musick (1999) "Application of Unified Evidence Accrual Methods to Robust SAR ATR," in I. Kadar (ed.), *Signal Processing, Sensor Fusion, and Target Recognition VIII*, SPIE Vol. 3720, pp. 71-79, ISBN 0-8194-3194-X

95. B.J. McCartin (1998) "Seven Deadly Sins of Numerical Computation," *American Mathematical Monthly*, December 1998, pp. 929-941

96. G. Mathéron (1975) *Random Sets and Integral Geometry*, J. Wiley

97. E. Michael (1951) "Topologies on Spaces of Subsets," *Trans. Am. Math. Soc.*, Vol. 71, pp. 152-182

98. R.A. Mitchell and J.J. Westerkamp (1999) "Robust Statistical Feature Based Aircraft Identification," *IEEE Trans. Aerospace & Electronic Systems*, Vol. 35 No. 3, pp. 1077-1094

99. E.W. Montroll (1952) "Markoff Chains, Wiener Integrals, and Quantum Theory," *Comm. on Pure & Applied Math.*, Vol. 5, pp. 415-453

100. S. Mori (2000) ``Random Sets in Data Fusion: Formalism to New Algorithms," *Proc. Third Int'l Conf. on Information Fusion*, Paris, July 10-13 2000

101. S. Mori (1998) "Multi-Target Tracking Theory in Random Set Formalism," *First Int'l Conf. on Multisource-Multisensor Information Fusion*, Las Vegas NV, July 1998

102. S. Mori (1998) "A Theory of Informational Exchanges--Random Set Formalism," *Proc. 1998 IRIS Nat'l Symp. on Sensor and Data Fusion*, Vol. I (Unclassified), Marietta GA, March 31-March 2, ERIM, pp. 147-158

103. S. Mori (1997) "Random Sets in Data Fusion: Multi-Object State-Estimation as a Foundation of Data Fusion Theory," in J. Goutsias, R.P.S. Mahler, and H.T. Nguyen (eds.), *Random Sets: Theory and Applications*, Springer-Verlag

104. S. Mori, C.-Y. Chong, E. Tse, and R.P. Wishner (1986) "Tracking and Classifying Multiple Targets Without A Priori Identification," *IEEE Trans. Auto. Contr.*, Vol. AC-31 No. 5, pp. 401-409

105. J.E. Moyal (1962) "The general theory of stochastic population processes," *Acta Mathematica*, Vol. 108, pp. 1-31

106. S. Musick, K. Kastella, and R. Mahler (1998) "A practical implementation of joint multitarget probabilities," in I. Kadar (ed.) *Signal Processing, Sensor Fusion, and Target Recognition VII*, SPIE Vol. 3374, pp. 26-37, ISBN 0-8194-2823-X

107. H.T. Nguyen (1978) ``On random sets and belief functions,`` *J. Math. Anal. and Appl.*, Vol. 65, pp. 531-542

108. A.I. Orlov (1977) ``Relationships between fuzzy and random sets: fuzzy tolerances,`` *Issledovania po Veroyatnostnostatishesk. Modelirovaniu Realnikh System*, Moscow

109. A.I. Orlov (1978) ``Fuzzy and random sets,`` *Prikladnoi Mnogomerni Statisticheskii Analys*, Moscow

110. L.I. Perlovsky (1997) "Cramér-Rao bound for tracking in clutter and tracking multiple objects," *Pattern Recognition Letters*, Vol. 18, pp. 283-288

111. N. Portenko, H. Salehi, and A. Skorokhod (1997) "On optimal filtering of multitarget tracking systems based on point processes observations," *Random Operators and Stochastic Equations*, Vol. 1, pp. 1-34

112. W.H. Press, S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery (1992) *Numerical Recipes in C: The Art of Scientific Computing*, Second Edition, Cambridge University Press

113. P. Quinio and T. Matsuyama (1991) ``Random Closed Sets: A Unified Approach to the Representation of Imprecision and Uncertainty,`` in R. Kruse and P. Siegel (eds.), *Symbolic and Quantitative Approaches to Uncertainty*, Springer-Verlag, pp. 282-286

114. C. Rago, M. Huff, B. Ravichandran, R.K. Mehra, T. Zajic, N. Lehtomaki, and R. Mahler (2001) "Optical path tracking in high-scintillation environments," in I. Kadar (ed.), *Signal Processing, Sensor Fusion, and Target Recognition X*, SPIE Vol. 4380, pp. 261-268, ISBN 0-8194-4075-2

115. B.D. Ripley (1976) "Locally finite random sets: foundations for point process theory," *Annals of Prob.*, Vol. 4 No. 6, pp. 983-994

116. L.H. Ryder (1996) *Quantum Field Theory*, Second Edition, Cambridge U. Press

117. W. Schmaedeke and K. Kastella (1993) "Information based sensor management," *Proc. 1993 Symp. On Command & Control Research*, Nat'l Defense University, June 28-29 1993, pp. 131-143

118. G. Shafer and R. Logan (1987) "Implementing Dempster's Rule for Hierarchical Evidence," *Artificial Intelligence*, Vol. 33, pp. 271-298

119. G.E. Shilov and B.L. Gurevich (1966) *Integral, Measure, and Derivative: A Unified Approach*, Prentice-Hall

120. A. Skorokhod (1998) "Optimal filtering in target tracking in the presence of uniformly distributed errors and false targets," June 1 1998, 1998 GTRI/ONR Workshop on Target Tracking and Sensor Fusion, Georgia Tech Research Institute, Marietta GA

121. D.L. Snyder and M.I. Miller (1991) *Random Point Processes in Time and Space*, Second Edition, Springer

122. E. Sorensen, T. Brundage, and R. Mahler (2001) "None-of-the-above (NOTA) capability for INTELL-based NCTI," in I. Kadar (ed.), *Signal Processing, Sensor Fusion, and Target Recognition X*, SPIE Vol. 4380, pp. 281-287

123. H.W. Sorenson (1988) "Recursive Estimation for Nonlinear Dynamic Systems," in J.C. Spall (ed.), *Bayesian Analysis of Statistical Time Series and Dynamic Models*, Marcel Dekker

124. H.W. Sorenson and D.L. Alspach (1971) "Recursive Bayesian Estimation Using Gaussian Sums," *Automatica*, Vol. 7, pp. 465-479

125. A. Srivastava (1996) "Inferences on Transformation Groups Generating Patterns on Rigid Motions," Ph.D. Dissertation, Washington University (St. Louis), Dept. of Electrical Engineering, August 1996

126. A. Srivastava, M.I. Miller, and U. Grenander (1991) "Jump-Diffusion Processes for Object Tracking and Direction Finding," *Proc. 29th Allerton Conf. on Communication, Control, and Computing*, U. of Illinois Urbana, pp. 563-570

127. A. Srivastava, N. Cutaia, M.I. Miller, J.A. O'Sullivan, and D.L. Snyder (1992) "Multi-Target Narrowband Direction Finding and Tracking Using Motion Dynamics," *Proc. 30th Allerton Conf. on Communication, Control, and Computation*, Sept. 30-Oct. 2, 1992, Monticello IL, pp. 279-288

128. M.C. Stein and C.L. Winter (1993) "An Additive Theory of Probabilistic Evidence Accrual," Los Alamos National Laboratories Report LA-UR-93-3336

129. L.D. Stone, C.A. Barlow, and T.L. Corwin (1999) *Bayesian Multiple Target Tracking*, Artech House

130. L.D. Stone, M.V. Finn, and C.A. Barlow (1997) "Unified Data Fusion," submitted for journal publication (manuscript copy, dated May 22 1997, provided by L.D. Stone)

131. D. Stoyan, W.S. Kendall, and J. Meche (1995) *Stochastic Geometry and Its Applications*, Second Edition, John Wiley & Sons

132. J.C. Strikwerda (1989) *Finite Difference Schemes and Partial Differential Equations*, Chapman & Hall

133. K.M. Tao, R. Abileah, and J.D. Lawrence (2000) "Multiple-target tracking in dense noisy environments: a probabilistic mapping perspective," in O.E. Drummond (ed.), *Signal and Data Processing of Small Targets 2000*, SPIE Vol. 4048, pp. 474-485, ISBN 0-8194-3674-7

134. H.L. van Trees (1968) *Detection, Estimation, and Modulation Theory, Part I: Detection, Estimation, and Linear Modulation Theory*, John Wiley & Sons

135. A. Wald (1949) "Note on the Consistency of the Maximum Likelihood Estimator," Annals of Math. Stat., Vol. 20, 595-601

136. E.L. Wahspress (1966) *Iterative Solution of Elliptic Systems and Application to the Neutron Diffusion Equations of Reactor Physics*, Prentice-Hall

137. R.B. Washburn (1987) "A Random Point Process Approach to Multi-Object Tracking," *Proc. Amer. Contr. Conf.*, Vol. 3, June 10-12 1987, Minneapolis, pp. 1846-1852

138. R.L. Wheedon and A. Zygmund (1971) *Measure and Integral: An Introduction to Real Analysis*, Marcel Dekker

139. T.O. Wolff, C.L. Lutes, R. Mahler, and D. Fixsen (1991) "Standards, Metrics, Benchmarks, and Monte Carlo: Evaluating Multi-Sensor Fusion Systems," *Proc. 1991 Joint Service Data Fusion Symp.*, Vol. I (Unclassified), Naval Air Development Center, Warminster PA, 1992, pp. 394-446

140. T. Zajic and R. Mahler (1999) "Practical information-based data fusion performance evaluation," in I. Kadar (ed.), *Signal Processing, Sensor Fusion, and Target Recognition VIII*, Vol. 3720, pp. 93-103, ISBN 0-8194-3194-X

141. T. Zajic, J. Hoffman, and R. Mahler (2000) "Scientific performance metrics for data fusion: new results," in I. Kadar (ed.), *Signal Processing, Sensor Fusion, and Target Recognition IX*, SPIE Vol. 4052, pp. 172-182, ISBN 0-8194-3678-X

142. T. Zajic, M. Huff, B. Ravichandran, and M.J. Noviskey (2001) "Joint tracking, pose estimation, and target recognition using HRRR and track data: new results," in I. Kadar (ed.), *Signal Processing, Sensor Fusion, and Target Recognition X*, SPIE Vol. 4380, pp. 196-206, ISBN 0-8194-4075-2

143. J.C. Naylor and A.F.M. Smith (1982) "Applications of a method for the efficient computation of posterior distributions," *Appl. Statist.*, vol. 31 no. 3, pp. 214-225

144. Y. Kharin (1996) *Robustness in Statistical Pattern Recognition*, Kluwer Academic Publishers

APPENDIX 1: CONSEQUENCES OF FINITE-SET STATISTICS

Section A summarized what FISST "is." The following subsections summarize some fundamental consequences of FISST:

(7) true Bayes-optimal multitarget nonlinear filtering (1-1);
(8) joint multitarget detection, localization, and identification (1-2);
(9) unified multi-evidence, multi-source, multi-target information fusion (1-3);
(10)unified multisource-multitarget information theory, with multitarget Cramér-Rao bounds (1-4);
(11)sensor management based on unified multisource-multitarget control theory (1-5); and
(12)unified multisource-multitarget decision theory and ROC curves (1-6).

1-1: TRUE BAYES-OPTIMAL MULTITARGET NONLINEAR FILTERING

FISST allows single-target nonlinear filtering to be directly generalized to the multisensor-multitarget realm. After summarizing the recursive Bayes filter, we discuss its generalization to the multitarget case (1-1-1) and summarize earlier work in multitarget Bayes filtering (1-2-2).

Recall that the foundation for single-sensor, single-target detection, tracking, and identification is the *recursive Bayes nonlinear filtering equations* [6,32,38,123],

Equation 1: Single-Target Bayes Filter

$$f_{k+1|k}(\mathbf{x} \mid Z^k) = \int f_{k+1|k}(\mathbf{x} \mid \mathbf{u}) f_{k|k}(\mathbf{u} \mid Z^k) d\mathbf{u}$$

$$f_{k+1|k+1}(\mathbf{x} \mid Z^{k+1}) = \frac{f_{k+1}(\mathbf{z}_{k+1} \mid \mathbf{x}) f_{k|k}(\mathbf{x} \mid Z^k)}{f_{k+1}(\mathbf{z}_{k+1} \mid Z^k)}$$

Equation 2: Single-Target Bayes-Optimal Estimators

$$\hat{\mathbf{x}}_{k|k}^{MAP} = \arg\sup_{\mathbf{x}} f_{k|k}(\mathbf{x} \mid Z^k)$$

$$\hat{\mathbf{x}}_{k|k}^{EAP} = \int \mathbf{x} \cdot f_{k|k}(\mathbf{x} \mid Z^k) d\mathbf{x}$$

where: $\mathbf{x}$ is the unknown state variable; $Z^k = \{\mathbf{z}_1,..., \mathbf{z}_k\}$ is the time-series of collected observations at time-step k; $f_k(\mathbf{z}|\mathbf{x})$ is the likelihood function; $f_{k+1|k}(\mathbf{y}|\mathbf{x})$ is the Markov transition density; $f_{k|k}(\mathbf{x}|Z^k)$ is the posterior distribution at time-step k; $f_{k+1|k}(\mathbf{x}|Z^k)$ is the prediction of this posterior to time-step $k+1$; $\hat{\mathbf{x}}_{k|k}^{MAP}, \hat{\mathbf{x}}_{k|k}^{EAP}$ are, respectively, the Bayes-optimal *a posteriori* (MAP) and expected *a posteriori* (EAP) estimates of the target state; and where

$$f_{k+1}(\mathbf{z}_{k+1} \mid Z^k) = \int f_{k+1}(\mathbf{z}_{k+1} \mid \mathbf{x}) f_{k|k}(\mathbf{x} \mid Z^k) d\mathbf{x}$$

is the Bayes normalization constant.

1-1-1 Multitarget Bayes Filtering. One might suspect that a straightforward, naïve generalization of the single-sensor, single-target Bayes filtering equations would lead to a similarly solid foundation for multisensor, multitarget information fusion. By this way of thinking, one would merely write down the following *multisensor-multitarget Bayes filtering equations* and declare victory:

Equation 3: Naïve Multitarget Bayes Filter

$$f_{k+1|k}(X \mid Z^{(k)}) = \int f_{k+1|k}(X \mid U) f_{k|k}(U \mid Z^{(k)}) \delta U$$

$$f_{k+1|k+1}(X \mid Z^{(k+1)}) = \frac{f_{k+1}(Z_{k+1} \mid X) f_{k|k}(X \mid Z^{(k)})}{f_{k+1}(Z_{k+1} \mid Z^{(k)})}$$

Equation 4: Naïve Multitarget Estimators

$$\hat{X}_{k|k}^{MAP} = \arg \sup_X f_{k|k}(X \mid Z^{(k)})$$

$$\hat{X}_{k|k}^{EAP} = \int X \cdot f_{k|k}(X \mid Z^{(k)}) \delta X$$

where: $X = \{x_1,\ldots, x_n\}$ is the unknown multitarget state-set; $Z^{(k)} = \{Z_1,\ldots,Z_k\}$ is the time-series of collected observation-sets at time-step k; $f_k(Z|X)$ is the likelihood function; $f_{k+1|k}(Y|X)$ is the multitarget Markov transition density; $f_{k|k}(X|Z^{(k)})$ is the multitarget posterior distribution at time-step k; $f_{k+1|k}(X|Z^{(k)})$ is the prediction of this posterior to time-step $k+1$; $\hat{X}_{k|k}^{MAP}, \hat{X}_{k|k}^{EAP}$ are, respectively, Bayes-optimal *a posteriori* (MAP) and expected *a posteriori* (EAP) estimates of the multitarget state-set; and where

$$f_{k+1}(Z_{k+1} \mid Z^{(k)}) = \int f_{k+1}(Z_{k+1} \mid X) f_{k|k}(X \mid Z^{(k)}) \delta X$$

is the Bayes normalization constant.

Surprisingly, the multitarget filtering equations cannot be copied from the single-target filtering equations in the blind fashion just indicated [57,60,62,67a,67b,70]. First, $\int \cdot \delta U$ and $\int \cdot \delta X$ are not conventional integrals but, rather, *set integrals* that account for the fact that numbers of both measurements and targets can vary. Second and as we shall see momentarily, the naïve "Bayes-optimal" multitarget MAP and EAP estimators $\hat{X}_{k|k}^{MAP}, \hat{X}_{k|k}^{EAP}$ of Equation 4 *do not even exist* [57,60,62,67a,67b,68,]. Third, there are a number of equally critical but more subtle errors implicit in the naïve approach. These issues are further explored in section 2-1-3 of Appendix 2 below.

1-1-2 Relationship With Earlier Approaches. Multitarget Bayes filtering is a relatively new concept in the multitarget information fusion community. If one assumes that the number of targets is known beforehand, the earliest exposition appears to be Washburn's [137] point-process formulation in 1987. When the number n of targets is *not* known and must be determined along with the individual target states, the earliest work appears to be due to Miller, O'Sullivan, Srivastava, et. al. [51,125,126,127]. Their very sophisticated approach utilizes solution of stochastic diffusion equations on non-Euclidean manifolds. It is also apparently the only approach to deal with continuous evolution of the multitarget state. The FISST approach was apparently the first to systematically deal with the general discrete state-evolution case (Bethel and Paras [8] assume discrete observation and state variables). Stone et. al. [130, pp. 161-207] have provided a valuable contribution by showing that the multi-hypothesis correlation (MHCT) tracker is a special case of the multitarget Bayes filter [24, p. 32], [62, p. 48]. Nevertheless their approach is best described as ``heuristic" for the reasons summarized in [62, pp. 41-42, 91-93], [66, pp. 222-223] and in section 2-2 of Appendix 2 below. Kastella's ``joint multitarget probabilities (JMP)" [40,43], introduced at LMTS in 1996, are nothing more than a renaming of a number of early core FISST concepts devised two years earlier (see [106, pp. 27-28] and section 2-2 of Appendix 2 below). These concepts include: set integrals, multitarget information metrics, multitarget posteriors, joint multitarget state estimators (joint multitarget detection, localization, and identification), etc. Portenko et. al., also using a point process approach, use branching processes to model target appearance and disappearance [111,120].

It should also be pointed out here that Mori, Chong, Tse and Wishner were the first to propose random set theory as a basis for multisensor-multitarget detection, tracking, and identification (albeit within a multi-hypothesis framework) **[104]**. Since 1995, Mori has returned to the field and produced a number of interesting random set-based papers **[100-103]**.

1-2: MULTITARGET ESTIMATION: JOINT MULTITARGET DETECTION, LOCALIZATION, AND IDENTIFICATION

Viewing a multitarget system as a single, unitary quantity leads to the concept of *simultaneous joint detection, localization, and identification*, first described as part of FISST in 1994 **[87,88,90]**. For example, let $f(\mathbf{z}|\mathbf{x})$ be a single-sensor, single-target likelihood function and let $\mathbf{z}_1,..., \mathbf{z}_m$ be data collected from the target. If data is conditionally independent, then the maximum likelihood estimate (MLE) of the target state $\mathbf{x}$ is

$$\hat{\mathbf{x}}^{MLE} = \arg\sup_{\mathbf{x}} f(\mathbf{z}_1 \mid \mathbf{x})\cdots f_k(\mathbf{z}_m \mid \mathbf{x})$$

The MLE is known to be optimal. It is also known to be identical to the Bayes-optimal maximum *a posterior* (MAP) estimate

$$\hat{\mathbf{x}}^{MAP} = \arg\sup_{\mathbf{x}} f(\mathbf{z}_1 \mid \mathbf{x})\cdots f_k(\mathbf{z}_m \mid \mathbf{x}) f_0(\mathbf{x})$$

when the prior distribution $f_0(\mathbf{x})$ is uniformly distributed. In like manner, let $f(Z|X)$ be the multitarget likelihood function for the same sensor and let $Z_1,...,Z_m$ be observation-sets collected from the targets. Then the following "global MLE"

$$\hat{X}^{MLE} = \arg\sup_{X} f(Z_1 \mid X)\cdots f_k(Z_m \mid X) = \arg\sup_{n,\mathbf{x}_1,...,\mathbf{x}_n} f(Z_1 \mid \{\mathbf{x}_1,...,\mathbf{x}_n\})\cdots f_k(Z_m \mid \{\mathbf{x}_1,...,\mathbf{x}_n\})$$

provides a *simultaneous, joint estimate of the most likely number $\hat{n}$ of targets, together with their most likely states $\hat{\mathbf{x}}_1,...,\hat{\mathbf{x}}_{\hat{n}}$ (position, velocity, identity) without resort to optimal report-to-track association.* Stated differently: the global MLE *optimally resolves the conflicting objectives of detection, identification, and localization.* Unlike the single-target case, and as is explained more fully in section 2-1-3 of Appendix 2 below, a multitarget analog of the MAP estimator *does not even exist.* Consequently, new multitarget Bayes estimators must be defined and their optimality must be demonstrated. LMTS has defined two such multitarget and shown them to be Bayes-optimal **[24, pp. 190-205]**, **[62, pp. 42-44]**:

- *Joint Multitarget Estimator (JoME):*

$$\hat{X}_{k|k}^{JoME} = \arg\sup_{X} f_{k|k}(X \mid Z^{(k)})\cdot \frac{c^n}{n!} = \arg\sup_{n,\mathbf{x}_1,...,\mathbf{x}_n} f_{k|k}(\{\mathbf{x}_1,...,\mathbf{x}_n\} \mid Z^{(k)})\cdot \frac{c^n}{n!}$$

- *Marginal Multitarget Estimator (MaME):*

$$\hat{X}_{k|k}^{MaME} = \arg\sup_{n} f_{k|k}(n \mid Z^{(k)})$$

$$f_{k|k}(n \mid Z^{(k)}) = \frac{1}{n!}\int f_{k|k}(\{\mathbf{x}_1,...,\mathbf{x}_n\} \mid Z^{(k)})d\mathbf{x}_1\cdots d\mathbf{x}_n$$

Here, c is a fixed constant whose units have been chosen so that $f(X) = c^{-|X|}$ is a multitarget density function. Both the JoM and MaM estimators have been shown to be Bayes-optimal, but only the MaME has been shown to be statistically consistent (i.e., converges to the correct answer). Moreover, because computation of the marginal density $f_{k|k}(n|Z^{(k)})$ loses information, one would expect that the MaME will converge more slowly than the JoME (if at all). This suspicion has been confirmed in analytical studies of simple model problems **[67b, pp. 150-151]**. Simple examples have also demonstrated that, because of

the information loss resulting from the computation of the marginal, the MaME ignores the certainty in targets whereas the JoME does not. **[67b, p. 149-150], [62, pp. 43-44]** .

During the Phase I contract, LMTS has also developed techniques for the direct nonparametric estimation of multi-object density functions using a generalization of the Parzen kernel technique **[24, pp. 312-336]**, **[85]**.

1-3: UNIFIED MULTI-EVIDENCE, MULTI-SOURCE, MULTI-TARGET INFORMATION FUSION

In the previous section, we assumed that all sources were sensors collecting unambiguous forms of data. Suppose, on the other hand, that the data is "ambiguous" in the sense of sections A.2.5 and B.3. Then the multisource-multitarget filtering equations of the previous section can be extended to include such data. This is especially the case if the sources can be assumed to be statistically independent. In this case, one derives formulas for the multisensor likelihood function using the techniques of the previous section, treating ambiguous data as though it were non-ambiguous. In effect, one simply computes a multisource-multitarget likelihood function as if the data were non-ambiguous, and then substitutes the generalized likelihood functions for the ambiguous sources in place of the "placeholder" exact likelihood functions. In this way, all sources—ambiguous or otherwise—can be subsumed under the FISST umbrella.

1-4: UNIFIED MULTISOURCE-MULTITARGET INFORMATION THEORY

FISST provides a means of directly generalizing single-sensor, single-target information-theoretic techniques to multisensor-multitarget situations. Consider the single-sensor, single-target case first. Suppose that a Kalman tracking algorithm is tracking a single target, whose ground truth state is known at any time. The question arises: At any instant of time, how much information is the tracker producing about the target, compared to ground truth? One useful way of answering this question is to compute the Kullback-Leibler *cross-entropy* or *discrimination*:

$$K(g_k; f_{k|k}) = \int g_k(\mathbf{x}) \log\left(\frac{g_k(\mathbf{x})}{f_{k|k}(\mathbf{x})}\right) d\mathbf{x}$$

where $f_{k|k}(\mathbf{x})$ is the posterior probability density associated with the Kalman filter output; and where $g_k(\mathbf{x})$ is another probability distribution associated with ground truth $\mathbf{g}_k$:

$$f_{k|k}(\mathbf{x}) = N_{P_{k|k}}(\mathbf{x} - \mathbf{x}_{k|k})$$

$$g_k(\mathbf{x}) = \begin{cases} V^{-1} & (\mathbf{x} \in B(\mathbf{g}_k)) \\ 0 & (\mathbf{x} \notin B(\mathbf{g}_k)) \end{cases}$$

Since any target has an actual spatial extent $B(\mathbf{g}_k)$ with (hyper)volume V^I , it is unnecessary to specify its location any more precisely than this extent. It can be shown that:

$$K(g_k; f_{k|k}) \cong -\log(V \cdot f_{k|k}(\mathbf{g}_k)) = K_0 + \log \det P_{k|k} + \frac{1}{2}(\mathbf{g}_k - \mathbf{x}_{k|k})^T P_{k|k}^{-1}(\mathbf{g}_k - \mathbf{x}_{k|k})$$

where K_0 is some constant.

1-4-1 Multitarget Information Measures of Effectiveness (MoEs). In like manner, the FISST "almost-parallel worlds principle" (APWOP) allows one to directly define the analogous concept for a multitarget information fusion and tracking algorithm:

51

$$K(g_k; f_{k|k}) = \int g_k(X) \log \left(\frac{g_k(X)}{f_{k|k}(X)} \right) \delta X \cong \log \left(V^\tau f(\{\mathbf{g}_1, ..., \mathbf{g}_\tau\}) \right)$$

where $f_{k|k}(X)$ is the *multitarget posterior* density associated with the multitarget tracker output; where $g_k(X)$ is another probability distribution associated with the ground truth set $G = \{\mathbf{g}_1, ..., \mathbf{g}_\tau\}$; and where the integral is now a *set integral*. This quantity is a measure of the multitarget tracker's instantaneous, over-all competence in estimating the numbers and states of the targets, relative to *complete knowledge of ground truth*. In some cases—for example, multi-hypothesis multitarget information fusion algorithms—this information metric can be approximated directly using explicit formulas, because in such cases the multitarget density function can be approximated using closed-form formulas. Efficient means have been determined for computing these information MoEs.

Such information MoEs are easily generalized to the case when ground truth is only imperfectly known (e.g. only to within the Cramér-Rao bound for the sensor). If ground truth is not known we can still estimate performance by measuring the algorithm's instantaneous, over-all competence relative to *complete ignorance about ground truth*:

$$K(f_{k|k}; u) = \int f_{k|k}(X) \log \left(\frac{f_{k|k}(X)}{u(X)} \right) \delta X$$

where $u(X)$ is the multitarget analog of the uniform distribution. Finally, these measures can be extended to include *user-defined* concepts of information. See [33,24,83,140,141] for more details.

1-4-2 Multitarget Cramér-Rao Bounds. Let $f(\mathbf{z}|\mathbf{x})$ be the likelihood function for a single sensor observing a single target. The generalized Cramér-Rao bound states that for a biased estimator $\mathbf{T}$ with covariance $\text{cov}_{\mathbf{T},\mathbf{x}}(\mathbf{w})$, that the accuracy of $\mathbf{T}$ is limited by the inequality

$$\left\langle \mathbf{w}, \text{cov}_{\mathbf{T},\mathbf{x}}(\mathbf{w}) \right\rangle \cdot \left\langle \mathbf{v}, \mathbf{L}_\mathbf{x}(\mathbf{v}) \right\rangle \geq \left\langle \mathbf{w}, \frac{\partial}{\partial \mathbf{v}} E_\mathbf{x}[\mathbf{T}(\mathbf{Z})] \right\rangle^2$$

where the directional derivative $\partial / \partial \mathbf{v}$ is applied to $E_\mathbf{x}[\mathbf{T}(\mathbf{Z})]$ as a function of $\mathbf{x}$, and $\mathbf{L}_\mathbf{x}$ is defined by

$$\left\langle \mathbf{c}, \mathbf{L}_\mathbf{x}(\mathbf{w}) \right\rangle = \int \frac{\partial \log f}{\partial \mathbf{v}}(\mathbf{z}|\mathbf{x}) \frac{\partial \log f}{\partial \mathbf{w}}(\mathbf{z}|\mathbf{x}) f(\mathbf{z}|\mathbf{x}) d\mathbf{x}$$

This form of the Cramér-Rao bound can be generalized to estimators of *vector-valued* outputs of multisource-multitarget algorithms. Let $\mathbf{J}$ be a vector-valued, multisource-multitarget state estimator—i.e., a vector-valued function $\mathbf{J}_m = \mathbf{J}(\Sigma_1, ..., \Sigma_m)$ of the random multisource-multitarget measurements $\Sigma_1, ..., \Sigma_m$ (assumed to be independent and identically distributed with multitarget likelihood functions $f(Z|X)$). Let $E_X[\mathbf{J}_m]$ be the expected value and covariance of the random vector $\mathbf{J}_m$. Define $L_{X,\mathbf{x},m}$ to be the unique (and necessarily linear) function that satisfies the identity

$$\left\langle \mathbf{v}, L_{X,\mathbf{x},m}(\mathbf{w}) \right\rangle = E \left[\frac{\partial \log f_m}{\partial_\mathbf{x} \mathbf{v}} \frac{\partial \log f_m}{\partial_\mathbf{x} \mathbf{w}} \right]$$

where $f_m(Z_1, ..., Z_m | X) = f_m(Z_1 | X) \cdots f_m(Z_m | X)$ and where the directional derivative of a function $f(X)$ of a finite-set variable X, if it exists, is defined as:

$$\frac{\partial f}{\partial_\mathbf{x} \mathbf{v}}(X) = \begin{cases} \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \left(f((X - \{\mathbf{x}\}) \cup \{\mathbf{x} + \varepsilon \mathbf{v}\}) - f(X) \right) & (\mathbf{x} \in X) \\ 0 & (\mathbf{x} \notin X) \end{cases}$$

Then the Cramér-Rao bound for the multisource-multitarget vector-valued estimator $\mathbf{J}$ is

$$\left\langle \mathbf{v}, C_{\mathbf{J}_m, X} \right\rangle \cdot \left\langle \mathbf{w}, L_{X,\mathbf{x},m}(\mathbf{w}) \right\rangle \geq \left\langle \mathbf{v}, \frac{\partial}{\partial_\mathbf{x} \mathbf{w}} E[\mathbf{J}_m] \right\rangle^2$$

52

for all **v,w**. See **[83]** and **[24, pp. 209-215]** for more details. An alternative approach for defining multitarget Cramér-Rao bounds can be found in **[110]**.

1-5: SENSOR MANAGEMENT VIA UNIFIED MULTISENSOR-MULTITARGET CONTROL THEORY

An adaptively guided sensor such as a missile-tracking camera exemplifies the single-sensor, single-target sensor management problem. This is also a classical control-theory problem: The camera must continually predict the target location on the basis of the target-observations it collects, and use this information to continually choose camera azimuth, elevation, and focal length in such a manner that the camera Field of View (FoV) overlaps the predicted target position as much as possible. In FISST, one views the entire multitarget system as a single "meta-target" following some trajectory in an abstract state space, and the entire sensor suite as a single "meta-sensor" which must be redirected in order to anticipate the predicted position of this meta-target. Given this, it is clear that the proper formulation of multisensor-multitarget sensor (and platform) management must be in terms of control theory. We briefly summarize a Bayes formulation of the single-sensor, single-target control problem (section 1-5-1), and then show how it can be directly generalized to the multisensor-multitarget case (see section 1-5-2).

1-5-1 Single-Sensor, Single-Target Sensor Management. In control theory, the sensor and target are analyzed as a *single joint system* rather than as two separate systems. The target has a state $\mathbf{x}$ (position, velocity, etc.), the sensor has a state $\mathbf{x}^*$ (azimuth, elevation, focal length, etc.), and the joint state is $\widetilde{\mathbf{x}} = (\mathbf{x}, \mathbf{x}^*)$. While the sensor collects observations $\mathbf{z}$ of the target state, it is itself being observed—by internal sensors that collect observations $\mathbf{z}^*$ of its state. So, observations of the joint state have the form $\widetilde{\mathbf{z}} = (\mathbf{z}, \mathbf{z}^*)$. The sensor is redirected by actuator mechanisms that (like the target) have physical and other limitations, e.g. slew rate. The behavior of these actuators is determined by control parameters $\mathbf{u}$ (voltages, etc.). One assumes a system-level measurement model for target and sensor

$$\widetilde{\mathbf{Z}}_{k+1} = (\mathbf{Z}_k, \mathbf{Z}_k^*) = (\mathbf{h}_k(\mathbf{x}, \mathbf{x}^*), \mathbf{h}_k^*(\mathbf{x}^*)) + (\mathbf{W}_k, \mathbf{W}_k^*) = \widetilde{\mathbf{h}}_k(\widetilde{\mathbf{x}}) + \widetilde{\mathbf{W}}_k$$

(That is, observations of the target depend on target state and sensor state, but observations of the sensor do not depend on target state.) Likewise, we assume a system Markov motion model

$$\widetilde{\mathbf{X}}_{k+1} = (\mathbf{X}_{k+1}, \mathbf{X}_{k+1}^*) = (\mathbf{g}_k(\mathbf{x}_k), \mathbf{g}_k^*(\mathbf{x}_k, \mathbf{x}_k^*, \mathbf{u}_k)) + (\mathbf{V}_k, \mathbf{V}_k^*) = \widetilde{\mathbf{g}}_k(\widetilde{\mathbf{x}}_k, \mathbf{u}_k) + \widetilde{\mathbf{V}}_k$$

(In other words, future target state depends only on current target state, whereas future sensor state depends on current target and sensor state, and the current control.) These equations can be reformulated in Bayesian fashion as a *system likelihood function* and *system Markov transition density*:

$$f_k(\mathbf{z}, \mathbf{z}^* \mid \mathbf{x}, \mathbf{x}^*) = f_k(\mathbf{z} \mid \mathbf{x}, \mathbf{x}^*, \mathbf{u}) f_k(\mathbf{z}^* \mid \mathbf{x}^*) = f_{\mathbf{W}_k}(\mathbf{z} - \mathbf{h}_k(\mathbf{x}, \mathbf{x}^*)) f_{\mathbf{W}_k^*}(\mathbf{z}^* - \mathbf{h}_k^*(\mathbf{x}^*))$$

$$f_{k+1|k}(\mathbf{y}, \mathbf{y}^* \mid \mathbf{x}, \mathbf{x}^*, \mathbf{u}) = f_{k+1|k}(\mathbf{y} \mid \mathbf{x}) f_{k+1|k}(\mathbf{y}^* \mid \mathbf{x}, \mathbf{x}^*, \mathbf{u}) = f_{\mathbf{V}_k}(\mathbf{y} - \mathbf{g}_k(\mathbf{x})) f_{\mathbf{V}_k^*}(\mathbf{y}^* - \mathbf{g}_k^*(\mathbf{x}, \mathbf{x}^*, \mathbf{u}))$$

Given streams $Z^k = \{\widetilde{\mathbf{z}}_1, ..., \widetilde{\mathbf{z}}_k\}$ and $U^k = \{\mathbf{u}_0, ..., \mathbf{u}_k\}$ of system data and control inputs, the conventional recursive Bayes filter (equation 1 of section 1-1) are used to propagate the system posterior $f_{k|k}(\widetilde{\mathbf{x}} \mid Z^k, U^{k-1})$. To provide a basis for closed-loop control, we assume that the Field of View is defined by a known probability of detection of the form $0 < p(\widetilde{\mathbf{x}}) \leq 1$ that tells us how probable it is that a target with state $\mathbf{x}$ will be observed by a sensor that is in state $\mathbf{x}^*$. In this case, the goal of the control system is to determine the control sequence U^k so that the expected probability of detection

$$\overline{p}(\mathbf{u}_k) = \int p(\widetilde{\mathbf{x}}) \cdot f_{k|k}(\widetilde{\mathbf{x}} \mid Z^k, U^{k-1}, \mathbf{u}_k) d\widetilde{\mathbf{x}}$$

is maximized over some time-range. For example, one can choose

$$p(\mathbf{x},\mathbf{x}^*) = \exp\left(-\frac{1}{2}(C\mathbf{x}-C^*\mathbf{x}^*)^T R^{-1}(C\mathbf{x}-C^*\mathbf{x}^*)\right)$$

where the positive-definite matrix S defines a Gaussian Field of View, C, C^* are matrices, and $C\mathbf{x}$ and $C^*\mathbf{x}^*$ are the *reference vector* and *controlled vector*, respectively.

1-5-2 Multisensor, Multitarget Sensor Management. The multisensor-multitarget control problem can be formulated in an analogous manner. We begin by concatenating all sensor state variables into a single vector $\vec{\mathbf{x}}^*$, all sensor measurement vectors into $\vec{\mathbf{z}}^*$, and all control vectors into $\vec{\mathbf{u}}$. In this case, the multisensor-multitarget state is $(X,\vec{\mathbf{x}}^*)$ and system-level observations are $(Z,\vec{\mathbf{z}}^*)$. Measurement and motion models for the complete multisensor-multitarget system have the form

$$f_k(Z,\vec{\mathbf{z}}^* \mid X,\vec{\mathbf{x}}^*) = f_k(Z \mid X,\vec{\mathbf{x}}^*)f_k(\vec{\mathbf{z}}^* \mid \vec{\mathbf{x}}^*)$$

$$f_{k+1|k}(Y,\vec{\mathbf{y}}^* \mid X,\vec{\mathbf{x}}^*,\mathbf{u}) = f_{k+1|k}(Y \mid X)f_{k+1|k}(\vec{\mathbf{y}}^* \mid X,\vec{\mathbf{x}}^*,\vec{\mathbf{u}})$$

Here, $f_k(Z \mid X,\vec{\mathbf{x}}^*)$ is the FISST multisensor-multitarget likelihood function; $f_k(\vec{\mathbf{z}}^* \mid \vec{\mathbf{x}}^*)$ is the multisensor-multiactuator likelihood function; $f_{k+1|k}(Y \mid X)$ is the FISST multitarget Markov density; and $f_{k+1|k}(\vec{\mathbf{y}}^* \mid X,\vec{\mathbf{x}}^*,\vec{\mathbf{u}})$ is the Markov transition density for the multisensor system. Given sequences $Z^{(k)} = \{(Z_1,\vec{\mathbf{z}}_1^*),...,(Z_k,\vec{\mathbf{z}}_k^*)\}$ and $U^k = \{\vec{\mathbf{u}}_0,...,\vec{\mathbf{u}}_k\}$ of system observations and multisensor controls, analogs of the multisensor-multitarget Bayes filter equations (Equation 3 of section 1-1) can be used to propagate the system multisensor-multitarget posterior $f_{k|k}(X,\vec{\mathbf{x}}^* \mid Z^{(k)},U^{k-1})$.

Unfortunately, the single-sensor, single-target control-theoretic reasoning used above cannot be transferred directly to the multisensor-multitarget case. This is because the probabilities of detection for all of the sensors are already included in the multisensor-multitarget likelihood function $f_k(Z \mid X,\vec{\mathbf{x}}^*)$. Instead, one must use the general reasoning employed by LMTS during the Phase I contract [68,77]: constructing density-level analogs of reference and controlled quantities. First, let $\vec{\mathbf{u}}_k$ be the current unknown control. Begin by predicting the current system posterior to the next time-step

$$\breve{f}_{k+1|k}(X,\vec{\mathbf{x}}^* \mid \vec{\mathbf{u}}_k) = \int f_{k+1|k}(X,\vec{\mathbf{x}}^* \mid W,\vec{\mathbf{w}}^*,\vec{\mathbf{u}}_k)f_{k|k}(W,\vec{\mathbf{w}}^* \mid Z^{(k)},U^{k-1})\delta W d\vec{\mathbf{w}}^*$$

Second, compute the system posterior, conditioned on a generic as-yet-uncollected multitarget observation Z^{k+1}:

$$\breve{f}_{k+1|k+1}(X,\vec{\mathbf{x}}^* \mid Z_{k+1},\mathbf{u}_k) \propto f_k(Z_{k+1} \mid X,\vec{\mathbf{x}}^*)f_{k+1|k}(X,\vec{\mathbf{x}} \mid Z^{(k)},U^{k-1},\mathbf{u}_k)$$

where the Bayes normalization constant is $f_k(Z_{k+1} \mid Z^{(k)},U^{k-1},\mathbf{u}_k)$. Our goal is to increase the informativeness of this "controlled" distribution—compared to the "reference" distribution $\breve{f}_{k+1|k}(X,\vec{\mathbf{x}}^* \mid \vec{\mathbf{u}}_k)$—in a manner that hedges against the fact that we do not know what actual system observation we will end up collecting. This amounts to the same thing as increasing the "peakiness" (decreasing the spread) of $\breve{f}_{k+1|k+1}$ compared to $\breve{f}_{k+1|k}$. This in turn is the same thing as increasing the overall peakiness of the ratio-function

$$r_{Z_{k+1},\vec{\mathbf{z}}_{k+1}^*}(X,\vec{\mathbf{x}}^* \mid \mathbf{u}_k) = \frac{\breve{f}_{k+1|k+1}(X,\vec{\mathbf{x}}^* \mid Z_{k+1},\mathbf{u}_k)}{\breve{f}_{k+1|k}(X,\vec{\mathbf{x}}^* \mid \mathbf{u}_k)} = \frac{f_k(Z_{k+1} \mid X,\vec{\mathbf{x}}^*)}{f_k(Z_{k+1} \mid Z^{(k)},U^{k-1},\mathbf{u}_k)}$$

as a function of $X,\vec{\mathbf{x}}^*$. One can do this by choosing some measure μ of peakiness and then averaging out both the multitarget states and the multitarget observations:

$$\mu(\mathbf{u}_k) = \int \mu \Big(r_{Z_{k+1}, \vec{z}_{k+1}^*} (X, \vec{\mathbf{x}}^* \mid \mathbf{u}_k) \Big) f(Z_{k+1} \mid Z^{(k)}) \delta Z_{k+1} \delta X d\vec{\mathbf{x}}^*$$

In our previous work, we chose an information theory-based measure, $\mu(x) = \log(x)$, and simplified the problem by choosing the "non-informative" null observation $Z_{k+1} = \varnothing$ instead of averaging over all observations. See section B.1 for additional details.

1-6: UNIFIED MULTISOURCE-MULTITARGET DECISION THEORY

The basic elements of decision theory can be directly generalized to the multisource-multitarget case. Suppose that we are trying to decide between two hypotheses H_0 and H_1 –for example, that a single target is present (hypothesis H_1) or not present (hypothesis H_0) in a cluttered scene. Suppose that $f(Z|X)$ is the multitarget likelihood for the sensor, where by assumption either $X = \varnothing$ or $X = \{\mathbf{x}\}$. Given a list $Z_1, \ldots, Z_m$ of cluttered observations collected from the scene, we are to decide between two possibilities: no target is present, or a target is present with state $\mathbf{x}_0$. If we form the likelihood ratio

$$R(Z_1, \ldots, Z_m) = \frac{f(Z_1 \mid \mathbf{x}) \cdots f(Z_m \mid \mathbf{x})}{f(Z_1 \mid \varnothing) \cdots f(Z_m \mid \varnothing)}$$

then, using set integrals, we can define a parametrized Receiver Operating Characteristic (ROC) curve:

$$p_{FA}(\tau) = \int_{L(Z_1, \ldots, Z_m) > \tau} f(Z_1 \mid \varnothing) \cdots f(Z_m \mid \varnothing) \delta Z_1 \cdots \delta Z_m$$

$$p_D(\tau) = 1 - \int_{L(Z_1, \ldots, Z_m) < \tau} f(Z_1 \mid \mathbf{x}) \cdots f(Z_m \mid \mathbf{x}) \delta Z_1 \cdots \delta Z_m$$

As with a conventional ROC curve, the slope of this ROC curve at any point is the value of the threshold that is required to achieve the probabilities of false alarm and detection corresponding to that point. See [72] for more detail.

APPENDIX 2: CRITICISMS OF FINITE-SET STATISTICS

The purposed of this Appendix is to describe and address certain criticisms of FISST that have been published during the duration of this contract. In section 2-1 I address these criticisms in a general fashion. Using this section as supporting material, I respond to the specific published criticisms in section 2-2.

2-1: GENERAL DISCUSSION OF THE CRITICISMS

FISST has attracted a great deal of positive attention in the information fusion and tracking communities. Such criticism as there has been has originated, oddly enough, with some researchers who (on the one hand) advocate the Bayesian approach because of its solid and systematic statistical foundations and in particular its "Bayes-optimality"; but (on the other hand) have attacked FISST by—in effect—arguing that solid and systematic statistical foundations are irrelevant in multitarget tracking! We will describe and respond to the specific published criticisms in section 2-2 below. First, however, it is necessary to address them in a systematic and general manner. When stripped of all circumlocution and condescension, these attacks reduce to the following single statement:

> *The multisource-multitarget engineering problems addressed by FISST actually require nothing more complicated than Bayes' rule; which means that FISST is of mere theoretical interest at best and, at worst, is nothing more than pointless mathematical obfuscation.*

This assertion is extraordinary less in the ignorance that it displays regarding FISST than in the ignorance that it displays regarding *Bayes' rule*. Two decades ago, J.C. Naylor and A.F.M. Smith noted that "The implementation of Bayesian inference procedures can be made to appear deceptively simple" **[143, p. 214]**. This is precisely what the critics of FISST have done. The seeming simplicity of Equations 1 and 2 of section 1-1 of Appendix 1 lulls many individuals into a failure to grasp the following fact: *both the optimality and the simplicity of the Bayesian framework can be taken for granted only within the confines of standard applications addressed by standard textbooks.* When one ventures out of these confines one must exercise proper engineering prudence—which includes verifying that textbook assumptions and presumptions still apply.

A major purpose of the monograph *An Introduction to Multisource-Multitarget Statistics and Its Applications* **[62]** and the book chapter *Random Set Theory for Target Tracking and Identification* **[60]** was to provide a detailed explanation of why this is the case. As I emphasized there, when one ventures away from standard applications addressed by standard textbooks and (unlike FISST) applies Bayesian approaches in a naïve or "cookbook" fashion, one can encounter severe difficulties. This is glaringly true in multi-object filtering. To understand these difficulties, we need to first review the critical assumptions that underlay the *ordinary* recursive Bayes filter (Equations 1 and 2 of section 1-1 of Appendix 1 above), and then show how a naïve generalization to the multisensor-multitarget case (equations 3 and 4 of section 1-1 of Appendix 1) ignores or glosses over these assumptions.

2-1-1 Single-Sensor, Single-Target Likelihood Functions. Bayes' rule *exploits to the best possible advantage the high-fidelity knowledge about the sensor contained in the likelihood function* $f_k(z|x)$. If $f_k(z|x)$ too imperfectly understood, then an algorithm will "waste" a certain amount N_{sens} of data trying (and perhaps failing) to overcome the mismatch between model and reality. Consequently, if we merely jot down Bayes' rule and declare victory we have either failed to understand that there is a potential problem or we are playing a shell game. If the former, we have failed to understand that our algorithm is *Bayes-optimal with respect to an imaginary sensor* unless we have the *true* likelihood function $f_k(z|x)$. If

the latter, we are avoiding the real algorithmic issue (what to do when likelihoods *cannot* be sufficiently well characterized) and instead implicitly pass the buck (the real issues and hard work) to the data simulation community.

In particular, how is it that we know that we have a *true* likelihood function? It is common practice in tracking and information fusion to model the random observation z_k produced by a sensor (without false alarms or missed detections) using a measurement model equation $Z_k = h_k(x) + W_k$ where W_k is a zero-mean random noise vector with density $f_{W_k}(z)$. Well-known textbooks tell us that the corresponding likelihood function is $f_k(z \mid x) = f_{W_k}(z - h_k(x))$. This likelihood is "true" in the sense that it is the the the *likelihood that actually corresponds to the measurement model.* However, how does one *explicitly construct* this true likelihood? We start with the *probability mass function* of the sensor model: $p_k(S|x) = \Pr(Z_k \in S)$. This is the total probability that the random observation Z_k will be found in any given region S if the target has state x. The probability mass $p_k(S|x)$ is just the sum of all the likelihoods in that region: $p_k(S|x) = \int_S f_k(z|x)dz$. From undergraduate calculus we know that

$$p_k(B_{\varepsilon,z} \mid x) = \int_{B_{\varepsilon,z}} f_k(z \mid x)dz \cong f_k(z \mid x)\lambda(B_{\varepsilon,z})$$

where $B_{\varepsilon,z}$ is some (hyper)ball of very small radius ε centered at z with (hyper)volume $V_{\varepsilon,z} = \lambda(B_{\varepsilon,z})$. Consequently, the following limiting ratio converges (given some mathematical complications we need not describe here) to the likelihood value:

$$f_k(z \mid x) = \lim_{\varepsilon \to 0} \frac{p_k(B_{\varepsilon,z} \mid x)}{\lambda(B_{\varepsilon,z})} = \lim_{\varepsilon \to 0} \frac{p_{W_k}(B_{\varepsilon,z-h_k(x)} \mid x)}{\lambda(B_{\varepsilon,z-h_k(x)})} = f_{W_k}(z - h_k(x))$$

as desired. This limiting ratio

$$\frac{\delta p_k}{\delta z} = \lim_{\varepsilon \to 0} \frac{p_{W_k}(B_{\varepsilon,z} \mid x)}{\lambda(B_{\varepsilon,z})}$$

is the *constructive Radon-Nikodým derivative* of the probability mass function $p_k(S|x)$ [119,138]. It provides an *explicit means of contructing* the (almost everywhere) unique probability density function $f_k(z|x)$ such that $p_k(S|x) = \int_S f_k(z|x)dz$ for all measurable subsets S. That is, it tells us how to construct the true likelihood function for a measurement model in the event that we cannot look it up in a textbook.

Finally, there is the kind of data—features extracted from signatures, English-language statements received over datalink, rules drawn from knowledge bases, etc.—that is so "ambiguous" (poorly understood from a statistical point of view) that probabilistic approaches in general—let alone the Bayes filtering equations—are not obviously applicable. Rather than seeing this as a gap in Bayesian inference that needs filling, a naïve viewpoint tends to sidestep the problem by ignoring it—and then all too frequently by condescending towards those who attempt to fill the gap with heuristic approaches such as fuzzy logic.

2-1-2 Single-Target Markov Densities. Much of what has been said about likelihoods $f_k(z|x)$ applies with equal force to Markov densities $f_{k+1|k}(y|x)$. The more accurately that $f_{k+1|k}(y|x)$ models target motion, the more effectively Bayes' rule will do its job. Otherwise, a certain amount N_{targ} of data must be expended in overcoming poor motion-model selection. The problem of constructing of the true Markov density from a motion model of the form $X_{k+1} = \Phi_k(x_k) + V_k$ is exactly analogous to that of constructing a true likelihood function from a measurement model.

2-1-3 Single-Target State Estimation. When we are faced with the problem of extracting an "answer" from the posterior distribution, complacency may encourage us to blindly copy state estimators from textbooks, or invent *ad hoc* ones. Great care must be exercised in the selection of a state estimator,

however. If it has unrecognized inefficiencies, then a certain amount N_{est} of data will be unnecessarily "wasted" in trying to overcome them—thought not necessarily with success. For example, the EAP estimator often produces erratic and inaccurate solutions when the posterior is multimodal (as occurs in applications with very low signal-to-noise ratio). Another example involves applications in which the state has the form $\mathbf{x} = (\mathbf{u}, \mathbf{v})$ where $\mathbf{u}$ involves kinematic state variables and $\mathbf{v}$ involves target-identity state variables. The joint MAP estimator is Bayes-optimal, convergent, etc. [134]:

$$(\hat{\mathbf{u}}^{JMAP}, \hat{\mathbf{v}}^{JMAP}) = (\mathbf{u}, \mathbf{v})^{MAP} = \arg\sup_{\mathbf{u},\mathbf{v}} f(\mathbf{u}, \mathbf{v} \mid Z^k)$$

However, we may be tempted to treat $\mathbf{u}$ as nuisance parameter and compute a joint estimate $(\hat{\mathbf{u}}^{MAP}, \hat{\mathbf{v}}^{MAP})$ using a marginal distribution:

$$\hat{\mathbf{u}}^{MAP} = \arg\sup_{\mathbf{u}} \int f_{k|k}(\mathbf{u}, \mathbf{v} \mid Z^k)d\mathbf{v}, \quad \hat{\mathbf{v}}^{MAP} = \arg\sup_{\mathbf{v}} f_{k|k}(\hat{\mathbf{u}}^{MAP}, \mathbf{v} \mid Z^k)$$

Because integration loses information about the state variable being regarded as a nuisance parameter, estimators of this type can converge more slowly than the joint MAP estimator. They will also produce noisy, unstable solutions when $\mathbf{u},\mathbf{v}$ are correlated and the signal-to-noise ratio is not large [67b, pp.148-149]. Last but not least, because a joint estimate $(\hat{\mathbf{u}}^{MAP}, \hat{\mathbf{v}}^{MAP})$ constructed in this manner is not a classical estimator, it is hardly clear that it is Bayes-optimal or convergent.

2-1-4 Computability of the Single-Sensor, Single-Target Bayes Filter. In general nonlinear problems, the integrals used for the predicted posterior $f_{k+1|k}(\mathbf{x}|Z^k)$ and the Bayes normalization constant $f_{k+1}(\mathbf{z}_{k+1}|Z^k)$ must be computed using numerical integration, and—since an infinite number of parameters are required to characterize the evolving posterior $f_{k|k}(\mathbf{x}|Z^k)$ in general—approximation is unavoidable [2,7,37,124]. Computational nonlinear filtering has become an area of active research in recent years [37,50]. LMTS has sponsored some of this research itself: particle-systems filters [3,46], Kouritzin's infinite-dimensional exact filter [47,48] (which has been applied to air traffic control problems [44]), and the spectral-separation techniques of Lototsky-Rozovskii [55]. A naïve viewpoint, by way of contrast, may tempt us to apply (deceptively) easy-to-understand textbook techniques that seem to promise high computational efficiencies. Naïve approximations, however, create the same difficulties as model-mismatch problems. An algorithm must "waste" a certain amount N_{appx} of data overcoming—or failing to overcome—accumulation of approximation error, numerical instability, etc. One example is the use of *central finite-difference schemes* to solve the Fokker-Planck equation for $f_{k+1|k}(\mathbf{x}|Z^k)$. In filtering problems the convection term of the FPE dominates the diffusion term. Under such circumstances, central differencing results in loss of probability mass (as well as the creation of negative probability) not only at the boundaries but throughout the region of interest, often resulting in poor solutions and numerical instability. This fact has long been known in the computational fluid dynamics community [19, p. 296]. Not only has this problem been cited as one of "Seven Deadly Sins of Numerical Computation" [95], it is so well known an error that it is cited as such in *Numerical Recipes in C* [112, p. 840].

One might also be tempted to argue that, in practical application, these difficulties can be overcome by simple *brute force*—i.e., by assuming that the data rate is high enough to permit a large number of computational cycles per unit time. In this case—or so the argument would go—the algorithm will function successfully despite its internal inefficiencies, because the total amount N_{data} of data that is collected is much larger than the amount amount $N_{ineffic} = N_{sens} + N_{targ} + N_{est} + N_{appx}$ of data required to overcome these inefficiencies. If this were the case, there would be few problems left to solve: most current challenging problems are challenging *either because data rates are not sufficiently high or because brute force computation cannot be accomplished in real time*. Consequently, in such situations brute force computation means the same thing as *non-realtime operation*.

These dangers become glaringly apparent in multitarget problems. Specifically:

2-1-5 Multisensor-Multitarget Likelihood Functions. Even if the single-sensor, single-target likelihood function $f_k(\mathbf{z}|\mathbf{x})$ can be determined with sufficient fidelity, what does one do in *multitarget problems*? We will "waste" data—or worse—unless we find the corresponding *true multitarget likelihood*—i.e., the *specific* function $f_k(Z|X) = f_k(\mathbf{z}_1,...,\mathbf{z}_m|\mathbf{x}_1,...,\mathbf{x}_n)$ that describes, with the *same* high fidelity as $f_k(\mathbf{z}|\mathbf{x})$, how likely it is that the sensor will collect observations $\mathbf{z}_1,...,\mathbf{z}_m$ (*m* random) given the presence of targets with states $\mathbf{x}_1,...,\mathbf{x}_n$ (*n* also random). Once again, if we are complacent we either fail to grasp that there is a problem—that our boast of "Bayes-optimality" is hollow unless we can construct the *provably* true $f_k(\mathbf{z}_1,...,\mathbf{z}_m|\mathbf{x}_1,...,\mathbf{x}_n)$—or we are encouraged to play another shell game. Which is to say, we construct a heuristic multitarget likelihood and *unwittingly or implicitly assume* that it is the true one.   To construct the true multisensor-multitarget likelihood function, at minimum we require *a multitarget generalization of the constructive Radon-Nikodým derivative* described in section 2-1-1. FISST addresses such issues by showing that familiar single-sensor, single-target reasoning—constructing measurement models and then constructing likelihood functions from them using such a derivative—can be directly generalized to the multitarget realm.

2-1-6 Multitarget Markov Densities. What does one do in the multitarget situation if the single-target Markov density $f_{k+1|k}(\mathbf{y}|\mathbf{x})$ is truly accurate? We must find the *provably true multitarget Markov transition density*—i.e., the *specific* function $f_{k+1|k}(Y|X) = f_{k+1|k}(\mathbf{y}_1,...,\mathbf{y}_r|\mathbf{x}_1,...,\mathbf{x}_n)$ that describes, with the *same* high fidelity as $f_{k+1|k}(\mathbf{y}|\mathbf{x})$, how likely it is that a group of targets that previously were in states $\mathbf{x}_1,...,\mathbf{x}_n$ (*n* random) will now be found in states $\mathbf{y}_1,...,\mathbf{y}_r$ (*r* also random)? Complacency may encourage us to simply assume that $f_{k+1|k}(\mathbf{y}_1,...,\mathbf{y}_n|\mathbf{x}_1,...,\mathbf{x}_n) = f_{k+1|k}(\mathbf{y}_1|\mathbf{x}_1)\cdots f_{k+1|k}(\mathbf{y}_n|\mathbf{x}_n)$—and then declare victory—meaning in particular that the number of targets is constant and target motions are uncorrelated. However, in real-world scenarios targets can appear (e.g., MIRVs and decoys emerging from a ballistic missile re-entry vehicle) or disappear (e.g., aircraft that drop beneath radar coverage) in correlated ways. Consequently, multitarget filters that assume uncorrelated motion and/or constant target number may perform poorly against dynamic multitarget environments, for the same reason that single-target trackers that assume straight-line motion may perform poorly against maneuvering targets. In either case, data is "wasted" in trying to overcome—successfully or otherwise—the effects of motion-model mismatch. FISST addresses this challenge by showing that familiar single-sensor, single-target reasoning—constructing motion models and then constructing Markov densities from them—can be directly generalized to the multitarget realm.

2-1-7 Multitarget State Estimation. In the multitarget case, the dangers of taking state estimation for granted become even more acute than in the single-target case. As already noted (section 1-2 of Appendix 1), *the multitarget versions of the standard MAP and EAP estimators are not even defined, let alone provably optimal.* The following simple example shows why (see **[62, pp. 40-42]** for a full discussion). Suppose that the multitarget posterior density has the simple form

$$f(X) = \begin{cases} 1/2 & if \quad X = \varnothing \\ \dfrac{1}{2}N_{\sigma^2}(x-1) & if \quad X = \{\mathbf{x}\} \\ 0 & if \quad |X| \ge 2 \end{cases}$$

where the variance σ^2 has units km^2. To compute the classical MAP estimate we must find the state $X = \varnothing$ or $X = \{x\}$ that maximizes $f(X)$. Since $f(\varnothing) = \frac{1}{2}$ is a unitless probability and $f(\{x\}) = \dfrac{1}{2}N_{\sigma^2}(x-1)$ has units of $1/km^2$, a naïve classical MAP asks us to compare the values of two quantities that are incommensurable because of *mismatch of units*. As a result, by simply changing units of measurement we can arbitrarily increase or decrease the numerical value of $f(\{1\})$—thereby getting

MAP estimates $X = \varnothing$ (there is no target in the scene) or $X = \{1\}$ (there is a target in the scene)! The posterior expectation estimate also fails. If it existed it would be

$$\int X \cdot f(X) \delta X = \varnothing \cdot f(\varnothing) + \int f(\{x\}) dx = \varnothing \cdot \frac{1}{2} f(\varnothing) + \frac{1}{2} \int N_{\sigma^2} (x-1) dx = \frac{1}{2} (\varnothing + 1km)$$

Once again, we have a *units-mismatch problem*: we are asked to add the unitless quantity $\varnothing$ to the united quantity *1 km*. Even if we assume that the continuous variable x is discrete, so that this problem disappears, we still must add the quantity $\varnothing$ to the quantity 1. If $\varnothing + 1 = \varnothing$ then $1 = 0$, which is impossible. If $\varnothing + 1 = 1$ then $\varnothing = 0$, so the same mathematical symbol represents two different states (the no-target state $\varnothing$ and the single-target state $x = 0$). The same problem occurs if we define $\varnothing + a = b_a$ for *any* real numbers a, b_a for then $\varnothing = b_a - a$.

One result is that *we must construct new multitarget state estimators and prove that they are well-behaved*. However, one common proof [135] that the MAP estimator will always converge to the correct answer requires the following assumptions: (i) the space of all measurements z is a topological space satisfying certain properties; (ii) the space of all states x is a metric space satisfying certain properties; and (iii) the likelihood $f_k(z|x)$ is measurable in the variable z (with respect to the measurement-space topology) and continuous in the variable x (with respect to the state-space metric). This means that words like "topology" and "measurable" can no longer be dismissively swept under the rug. One of LMTS's earliest accomplishments in the Phase I contract was to use FISST techniques to construct Bayes-optimal multitarget state estimators and show that they are well-behaved [24, pp. 190-205]. This work was summarized in section 1-2 of Appendix 1 above.

2-1-8 Common Errors Involving Multi-Object Integrals ("Set Integrals"). The careless assumption that single-target Bayes filtering can be generalized to the multitarget case in a "straightforward way" has led many researchers into a number of fundamental errors. The most common errors result *from a failure to notice that many kinds of single-target integrals cannot be directly generalized to multitarget integrals (set integrals) because of the units-mismatch problem just described in section 1-2 of Appendix 1*. For example, one author has assumed that the L_2 metric can be directly generalized to multitarget densities $f(X)$, $g(X)$:

$$\| f - g \|_2^2 = \int (f(X) - g(X))^2 \delta X$$

Another author has assumed that Shannon entropy can be likewise generalized:

$$\varepsilon_f = -\int f(X) \log f(X) \delta X$$

However, and as LMTS has repeatedly pointed out in a number of publications (see, for example, [24, pp. 163, 303], [62, p. 39]) *neither of these integrals are defined in general*. For example, in one dimension

$$\int (f(X) - g(X))^2 \delta X$$

$$= (f(\varnothing) - g(\varnothing))^2 + \int (f(\{x\}) - g(\{x\}))^2 dx + \frac{1}{2} \int (f(\{x_1, x_2\}) - g(\{x_1, x_2\}))^2 dx_1 dx_2 + \ldots$$

However, if x is a continuous variable with units in (say) meters, then sum is undefined because its first term is unitless, its second term has units of $1/m$, its third term has units of $1/m^2$, and so on. Consequently, the indicated sum is meaningless.

2-1-9 Computability of the Multisensor-Multitarget Bayes Filter. If the single-sensor, single-target Bayes filter is so computationally challenging that it must be approximated, then the multitarget nonlinear filtering equations will never be of practical interest without the development of drastic but intelligent approximation strategies. FISST provides systematic, principled approximation methods based on generalizations of well-known single-sensor, single-target approaches. Under our last USARO contract,

60

approximations based on a multitarget analog of the Gaussian approximation [64] and a statistical analog of the α-β-γ filter [16,58,59,61,65] were devised. These are summarized in section B.6.

2-2: RESPONSE TO PUBLISHED CRITICISMS

We now turn to the specific published criticisms mentioned at the beginning of this Appendix. These have taken the form of peremptorily sweeping dismissals of FISST from two sources: L. Stone and his associates; and K. Kastella. We will address each in turn.

In 1996, Stone and his associates introduced a multitarget tracking approach which they have variously called "unified data fusion" or "the likelihood approach" [5,130], [129, pp. 161-207] and which consists essentially of the naïve multitarget Bayes filter and the naïve multitarget estimators (equations 3 and 4 of section 1-1 of Appendix 1). Apparently unaware at the time of the existing literature in this area (as summarized in section 1-1-2 of Appendix 1), they have responded to it by trying to manufacture spurious distinctions and deficiencies. FISST and the "jump diffusion" work of Miller, O'Sullivan, Srivastava, et. al. have been special objects of attention. Regarding FISST, in a 1999 book Stone, Barlow, and Corwin made the following statements:

> "...Mahler develops an approach to tracking that relies on random sets. The random sets are composed of finite numbers of contacts so that this approach applies only to situations where there are distinguishable sensor responses that can clearly be called out as contacts or detections. In order to use random sets, one needs to specify a topology and a rather complex measure on the measurement space for the contacts. The approach...requires that the measurement spaces be identical for all sensors. In contrast, the likelihood function approach used in this book, which transforms sensor information into a function on the target state space, is simpler and appears to be more general...[and] allow[s] one to handle situations in which sensor responses are not strong enough to call contacts." [129, pp. 204-205]

As for Kastella, in 1996 he employed an approach he called "joint multitarget probabilities" or "JMP" to generalize a single-sensor, single-target sensor management techique called "discrimination gain" to the unknown-n multitarget case. During the years 1993-1998 Kastella was intimately familiar with FISST while employed at LMTS, and "JMP" is nothing more than a new name and notation for a special case of certain basic FISST concepts devised two years earlier. These concepts include: multi-object density functions, multitarget posteriors, set integrals, multitarget Kullback-Leibler MoEs, joint multitarget estimators, the almost-parallel worlds principle (APWOP), etc. This fact is specifically acknowledged in a 1998 paper that Kastella co-authored:

> "JMP, and the conceptual apparatus surrounding it, are elements of a comprehensive approach to data fusion (including multisensor-multitarget detection, tracking, classification, sensor management, multi-evidential fusion and performance estimation) called 'finite-set statistics' (FISST) FISST...was invented because (1) true Bayes-optimal multitarget estimation and filtering encounters fundamental theoretical and practical difficulties when the number of targets is unknown, and (2) these problems get worse when 'ambiguous' data (e.g. attributes, natural-language statements, rules) are present [106, p. 27]

As just one example, Kastella's generalization of "discrimination gain" is based on the specific application of the APWOP described at the end of section A.2.1—one which has been used as a standard example since 1994 [68, pp.256-258]:

$$K(f;g) = \int f(\mathbf{x}) \log\left(\frac{f(\mathbf{x})}{g(\mathbf{x})}\right) d\mathbf{x} \quad \overset{APWOP}{\Rightarrow} \quad K(f;g) = \int f(X) \log\left(\frac{f(X)}{g(X)}\right) \delta X$$

Now employed elsewhere, Kastella has, like Stone et. al., been laboring to distinguish "JMP" from FISST by manufacturing spurious distinctions and deficiencies—in a nutshell, that "JMP" is a great advance over FISST because it is supposedly vastly simpler. Kastella has written:

"...NLF [nonlinear filtering] is a particular example of Bayesian filtering that generalizes in a straightforward way to multitarget applications..." **[41, p. 256]**

"One way to characterize this collection of targets [the multitarget state $x_1,..., x_N$] at a particular time is to use Bayesian methods to construct the conditional probability density $p(x_1,..., x_N | Y)$ [the multitarget posterior]. This can be computed using standard Bayesian methods, requiring no fuzzy or random set concepts to be introduced..." **[42, p. 1]**

In addition, in an anonymous review last year of a third-party paper, Kastella or an associate wrote:

"Although JMP and FISST are attacking some of the same problems, there are important distinctions between them. JMP is a straightforward application of purely Bayesian concepts to the problem of search, track and identification, with the confounding issue that target count is unknown and must be estimated too. FISST on the other hand is a theoretical framework for unifying most techniques for reasoning under uncertainty (e.g. Dempster Shafer, fuzzy, Bayes, rules) in a common structure (based on random sets)...JMP is derived from first principles..., makes no appeal to random sets or related concepts,...[and] is a general mathematical technique for evidence accrual that is able to ingest detection, track, and identification evidence equally well."

In leveling such unequivocal and sweeping attacks in print—charging, in effect, that FISST is nothing more than obviously pointless mathematical obfuscation—all of these authors have, in the process, tacitly opened the door for equally vigorous scrutiny of their own technical claims. We have responded to Stone et. al. in the recent publications **[62, pp. 41-42, 91-93]**, **[66, pp. 222-223]**, and to Kastella in the publications **[67a,67b]**. What follows is a condensation and elaboration of those counter-criticisms. We show why: (1) the criticisms of FISST are false; (2) these authors' claims of simplicity are spurious and based on fundamental ignorance of the assumptions underlying the Bayesian approach; and (3) this ignorance leads them into error.

What all of these authors share in common is a failure to grasp the elementary points made in sections 2-1-1 through 2-1-9 of this Appendix. *It is easy to claim invention of a simple and yet elastically all-subsuming theory if—like these authors—one does so by dealing only with simple special cases and then avoiding, neglecting, or glossing over the technical specifics that would be required to actually substantiate their expansive extrapolations to the general case. Likewise, it is easy to portray earlier approaches as deficient if—as with these authors—one does so through misrepresentation and application of technical double standards.* Specifically:

(1) The approaches claimed by these authors are so imprecisely formulated that they have all found it possible to *both disparage and unwittingly assume "random set concepts" at the same time*;

(2) The true multitarget likelihood function $f_k(Z|X)$ and the true multitarget Markov density $f_{k+1|k}(Y|X)$ are useless mathematical abstractions, unless one has—as these authors do not—general and explicit procedures for constructing multitarget measurement models and multitarget motion models, and for constructing $f_k(Z|X)$ and $f_{k+1|k}(Y|X)$ from these models.

(3) A "Bayes-optimal multitarget filter" cannot be Bayes-optimal unless one has, as these authors do not an explicit approach for constructing: (a) the "true" Bayes posterior; (b) the provably true multitarget likelihood function that is required to construct the true Bayes posterior; and (c) a multitarget state estimator that is provably well-defined, Bayes-optimal, convergent, etc.

(4) The FISST approach recognizes the fact that one cannot blindly assume—as these authors do—that the classical Bayes-optimal estimators generalize in a "straightforward way" to the general multitarget case. It also recognizes—as these authors do not—that one must not only devise new multitarget state estimators, but show that they are well-defined, Bayes-optimal, convergent, etc.;

In the case of Stone et. al., we demonstrated that their specific criticisms of FISST are without foundation:

(1) FISST has *always* had explicit procedures for dealing with unknown numbers of targets [87,88,90];

(2) The FISST procedure for constructing multitarget likelihood functions *does* subsume those sensors whose observations are superpositions of signals from the individual targets [62, p. 18];

(3) FISST has *always* been capable of dealing with both post- and pre-detection measurements and with multiple sensors having different measurement spaces. Specifically, it is common practice to model pre-detection observations as vectors whose components are image pixel intensities, radar range-bin intensities, etc. Consequently, nearly any measurement space is a subspace of $\mathbb{R}^{m(s)} \times C_s \times \{s\}$ where s is a sensor tag and where $\mathbb{R}^{m(s)}$ and C_s denote continuous and discrete measurement variables, respectively [24, p. 220], [90, p. 32]. The multisensor-multitarget likelihood function $f(Z|X)$ is meaningless unless one first bundles all observations into a *single multisensor measurement space*, namely the topological sum (disjoint union) of the individual measurement spaces:

$$(\mathbb{R}^{m(1)} \times C_1 \times \{1\}) \; \dot{\cup} \; ... \; \dot{\cup} \; (\mathbb{R}^{m(e)} \times C_e \times \{e\})$$

In turn, this space is a subset of the product space $\mathbb{R}^M \times C$ where $M = m(1) + ... + m(e)$ and $C = (C_1 \dot{\cup} ... \dot{\cup} C_e) \times \{1,...,e\}$. To avoid unnecessary notational and mathematical complexity it is, therefore, sufficient (as in FISST) to use only $\mathbb{R}^M \times C$.

(4) Stone et. al. do not appear to understand that to define a *Bayesian* multisensor-multitarget likelihood $f(Z|X)$ one must (a) precisely define a *single* multisensor-multitarget observation-space in full generality; (b) "specify a topology" for this space; (c) define a random variable on this space in terms of this topology; and (d) define $f(Z|X)$ as a conditional probability density in terms of this random variable.

Furthermore, we demonstrated that the advertised simplicity and generality of their "likelihood approach" is spurious and leads them into errors, one of which is *implementationally fatal*. Specifically:

(1) Its theoretical basis is so imprecisely formulated that, on the single occasion that Stone et. al. have attempted to define multitarget observations with some degree of precision [130, pp. 5-6], they have unwittingly assumed a random set framework: "let Y_k be the set of values of sensor observations received at time t_k. Let y_k denote a value of the random variable Y_k. However, if Y_k is both a "random variable" and a "set" (of measurements collected by several sensors), then it is a *randomly varying finite subset*! Such a random variable *does not even make sense* unless first one defines—as in FISST—*some* "topology" and "measure" on the class of all finite subsets.

(2) Its Bayes-optimality and "explicit procedures" are both frequently asserted but never actually justified or even described with precision;

(3) Its claimed "general approach" for dealing with unknown numbers of targets is fatally flawed: "The [multitarget] posterior distribution...constitutes the Bayes estimate of the number and state of the targets...From this distribution we can compute other estimates when appropriate such as maximum a posteriori probability estimates or means" [129, pp. 162-163]. Contrary to this assertion, posteriors are not "estimators" of state variables; the multitarget MAP can be defined only when state space is discretized (or when all continuous variables are unitless); and a multitarget posterior expectation apparently cannot be defined at all (see sections 1-2 of Appendix 1 and 2-1-3 of this Appendix). This is especially ironic, given that Stone et. al. claim the superiority of "the likelihood approach" because it possesses the above (nonexistent) "explicit procedure" for dealing with an unknown number of targets, whereas earlier work supposedly does not;

(4) These authors subscribe to double standards: on the one hand, they repeatedly (and erroneously) lambast all earlier researchers for lacking "explicit methods' for various things; but on the other hand, they feel no need to specify—as in FISST—"explicit methods" for concepts that, without these methods, are *useless mathematical abstractions*—e.g. *general* multitarget likelihood functions,

general multitarget Markov models, *general* multitarget integrals, etc. Indeed, their very failure to specify such explicit methods is what allows them to portray the "likelihood approach" as simple.

(5) They also engage in "Catch-22's": When FISST is illustrated using specific examples (e.g., sensors with post-detection observations), Stone et. al. seize on these as proof that FISST is not general. But when FISST is developed in full generality, they seize upon this as proof that FISST is not simple.

(6) Even if "the likelihood approach" had been developed with precision and without error, it would still have been entirely subsumed by the earlier work of Miller, O'Sullivan, Srivastava, et. al., dating from 1991 (see section 1-1-2 of Appendix 1). Claims to the contrary withstanding **[129, pp. 204-205]**, Miller et. al. *do* provide an *explicit* procedure (not to mention a general and sophisticated algorithmic implementation) for dealing with moving and unknown numbers of targets. Miller et. al. also *explicitly* address many observation-types other than "camera images" and, in particular, sensors for which "the received signal is a superposition of the signals from each target" **[127, p. 284]**.

Unlike Stone et. al., Kastella has offered no specific substantiations for his own brusque dismissal of FISST. Instead, he has chosen only to cite the above statements of Stone et. al.—thereby unwittingly inheriting their deficiencies—and his first "JMP" paper **[43]**. We have responded to Kastella's more covert attacks with specific counter-arguments in the publications **[67a,67b]**. Specifically:

(1) Like Stone et. al., Kastella achieves a façade of all-encompassing simplicity by ignoring or glossing over basic issues. Specifically, he: (a) elastically extrapolates from special cases such as toy scenarios, discrete state spaces, single-observation sensor measurement models, etc.; (b) uses simplistic multitarget motion models (section 2-1-6); (b) defines multitarget likelihood functions but does not show why they are not *ad hoc* contrivances (section 2-1-5); (c) defines a multitarget estimator (the MaM estimator of section 1-2 of Appendix 1), but does not show that it is Bayes-optimal, convergent, etc.; and (d) glosses over the combinatoric difficulties involved in extending single-target computational approaches to the multitarget case (section 2-1-9);

(2) This spurious simplicity leads to outright error: "...[if] the target space is discretized into a collection of cells [then] in the continuous case, the cell probabilities can be replaced by densities in the usual way" **[40, p. 123]**. Contrary to this assertion, multitarget state filtering encounters basic difficulties when continuous variables are present (sections 1-2 of Appendix 1 and 2-1-2, 2-1-7, and 2-1-8 of this Appendix). At the very least, the naïve multitarget MAP estimator (Equation 4 of section 1-1 of Appendix 1) fails because it produces different answers when units of measurement are changed.

(3) Kastella is perfectly aware of the significance and seriousness of this kind of error, since he has reproached others for having committed it. His "discrimination gain" approach arose when, in 1993, he and a co-author noticed that earlier researchers' sensor management techniques suffered from exactly the same kind of deficiency. To wit: "Several authors...suggest using the Trace of the covariance matrix as the measure of information. There is a units problem in doing this. Hintz...attempts to resolve the units problem...[b]ut this change is not invariant under even a change of units." **[117, pp. 139-140]**

(4) To manufacture distinctions between "JMP" and FISST, Kastella has to resort to hair-splitting and misrepresentation. For example:

> "One technical difference...is that, as in the example above, JMP maintains separate densities for the 1-target and 2-target cases. On the other hand, random sets treat both the 1-target and 2-target cases with a single density, say $f(x_1, x_2)$. Then the random set density for two targets with one at x_1 and the other at x_2 is $f(x_1, x_2)$ while the density for a single target whose location is x is $f(x,x)$." **[40, pp. 123]**

FISST indeed uses a single density $f(X)$ to subsume all target numbers whereas "JMP" uses a separate density for each case. However, *a trivial change of notation does not constitute a "technical difference."* A "JMP" is just a FISST multitarget posterior distribution, first described in 1994 **[90, p. 337]**, assuming that state space is discrete: $f_{FISST}(\{x_1,..., x_n\}) = n! \, f_{"JMP"}(x_1,..., x_n)$. As Kastella is

prefectly aware, both notations have been employed interchangeably in FISST (though random set notation is to be preferred because of the advantages outlined in section A.2.1). Moreover, in FISST the single-target case is represented as $f(x) = f(\{x\}) = f(\{x,x\})$ and not by "$f(x,x)$".

Although Kastella is aware of these counter-criticisms, his only response has been to issue more overtly dismissive and authoritative-sounding—but equally unsubstantiated—pronunciamentos. Ironically, in the process he has only compounded his difficulties:

(1) Like Stone et. al., Kastella disparages "random set concepts" even while unwittingly assuming them. Immediately after explicitly proclaiming the pointlessness of such concepts, he tries to develop a multitarget Bayes filter for a "stochastic distribution" called a "multi-target microdensity," which he defines as $\rho = \delta_{\mathbf{x}_1} + ... + \delta_{\mathbf{x}_n}$ [42]. However, this is nothing more than *an unwitting re-invention of a simple point process* $\delta_X = \delta_{\mathbf{x}_1} + ... + \delta_{\mathbf{x}_n}$ *in random density form*—that is, *an unwitting re-invention of a random finite set* $X = \{\mathbf{x}_1, ..., \mathbf{x}_n\}$ *expressed in more complex notation* (see section B.6.2).

(2) Likewise, his so-called "probability density functional" on the microdensity, $p\{\rho \mid Y\} = p\{\delta_{\mathbf{x}_1} + ... + \delta_{\mathbf{x}_n} \mid Y\}$, is—like his so-called "JMP"—just a new name and notation for a FISST multitarget posterior density $f(X|Y) = f(\{\mathbf{x}_1, ..., \mathbf{x}_n\}|Y)$.

(3) The "microdensity" approach requires extremely complex functional analysis—in particular, complicated and very difficult-to-evaluate integrals $\int \cdot D\rho$ defined on function spaces [20,99]. Yet when restricted to ρ's that are "microdensities," these integrals turn out to be nothing more than FISST set integrals: $\int f\{\rho\} D\rho = \int f\{\delta_X\} \delta X = \int f(X) \delta X$.

(4) Consequently, his "Bayesian filter on the microdensity" amounts to nothing more than an elaborate (if ongoing) failure to recognize that *a change of name and notation does not add new substance*. For example, things will appear even more impressively complicated (but no substance will be added) if one rewrites a random set as $\delta_{\delta_{\mathbf{x}_1}} + ... + \delta_{\delta_{\mathbf{x}_n}}$ (i.e., as a sum of Dirac deltas concentrated on Dirac deltas), calls it a *multi-target picodensity*, and then assigns some inventive new name to the multitarget posterior $p\{\delta_{\delta_{\mathbf{x}_1}} + ... + \delta_{\delta_{\mathbf{x}_n}} \mid Y\}$.

(5) Similarly, Kastella succeeds only in unwittingly replicating other years-old FISST concepts in highly obfuscated form. What he calls the "expected value of the target density," $\bar{\rho} = \int \rho \cdot p\{\rho \mid Y\} D\rho$, is just the point process first-moment density, or *probability hypothesis density* $D(\mathbf{x})$, introduced as part of FISST in 1996 (see [24, pp. 168-169], [74, p. 149] and section B.6.2 below):
$$\bar{\rho}(\mathbf{x}) = \int \delta_X(\mathbf{x}) \cdot f(X \mid Y) \delta X = \int_{X \supseteq \{\mathbf{x}\}} f(X \mid Y) \delta X = D(\mathbf{x})$$

(6) Like Stone et. al., Kastella engages in double standards. "JMP" renders "random set concepts" irrelevant because, or so it is asserted, "JMP" *already* addresses the multitarget Bayes filter rigorously and completely in a "straightforward way" that requires no pointless mathematical complexity— specifically, *no nonlinear filtering based on simple point processes*. Given this, what possible need could there be (other than to manufacture pointless but awe-inspiring mathematical complexity) for an approach such as "classical Bayesian nonlinear filtering methods...extended to the microdensity"—i.e., for *nonlinear filtering extended to simple point processes*?! Stranger still, the added complexity is so great that (according to Kastella) his "new" approach is not yet rigorous. What then could be the point, given that all necessary rigor in multitarget filtering has purportedly been achieved years ago via the parsimoniously simple "JMP"?

(7) Given Kastella's highly authoritative declaration of the irrelevance of "fuzzy...set concepts" it is ironic that if one assumes (as he does) that state space is discrete, then his "expected value of the

65

target density" is nothing more than *a fuzzy subset of state space,* defined by the fuzzy membership function $\overline{p}(x) = \Pr(x \in X)$ where X denotes the random track-set **[24, p. 169]**, **[65, p. 108]**.

Last but not least, in eschewing specific substantiation Kastella implicitly adopts the following rhetorical stance: Such substantiation is unnecessary because the pointlessness of "random set concepts" should be as obvious to any technically competent individual as it is to him. In so doing, he has chosen to invoke his own technical authority as a central element of his substantiation. However, it is easy to show that Kastella's misunderstandings of Bayes multitarget filtering stem from a broader failure to grasp the dangers of a naïve, textbook perspective *even in the single-target case* (as summarized in sections 2-1-1, 2-1-2, and 2-1-3 of this Appendix). The truth of this can be seen in a recent, ostensibly authoritative book chapter devoted to Bayes nonlinear filtering based on a numerical Fokker-Planck equation (FPE) solver called the "alternating direction implicit (ADI)" **[136]** method:

(1) Kastella erects a façade of spurious simplicity by (a) citing a standard textbook result **[38, p. 165]** in Bayes filtering, namely that the predicted posterior is a solution of the FPE; (b) using another standard textbook approach (ADI) **[132, pp. 142-150]** to transform it into a long-standing unsolved problem, namely real-time numerical solution of PDEs using finite-difference methods; (c) declaring victory as an algorithm designer; and then (d) camouflaging this sleight-of-hand by redefining this problem as an "enabling technology that is key to making NLF practical" **[41, pp. 235-236]**.

(2) In addressing the specific application of HRRR joint tracking and identification, Kastella conjures up a similarly deceptive simplicity. By employing his textbook filter in conjunction with a toy HRRR signature model that sidesteps serious thinking about the real engineering issues **[41, p. 253]**, he again declares victory and then shunts responsibility for these issues off onto yet another "enabling technology." Specifically, such an algorithm will be useless for practical application without the ultra-high-fidelity HRRR simulation algorithms that many HRRR experts **[98]** doubt can ever be implemented in real time (see section 2-1-1).

(3) Instead of using a known Bayes-optimal joint state estimator to accomplish joint tracking and identification, Kastella chooses the *ad hoc* joint estimator described in section 2-1-3 **[41, p. 252]**. In so doing, he appears oblivious to the fact that his approach is "Bayesian" in name only, if it glosses over fundamental issues such as Bayes risk, convergence, computational inefficiency, etc.;

(4) After authoritatively warning readers to beware the pitfalls of numerically unstable explicit finite-difference methods **[41, p. 235]**, Kastella himself then—with equal authoritativeness—tumbles into exactly the same pit when he employs central differencing **[41, p. 243]** (the numerically unstable "deadly sin of numerical computation" described above in section 2-1-4).

Items (1) and (2) in particular should be contrasted with the perspective taken by credentialed experts in computational nonlinear filtering. All have assumed that, to be serious in this line of research, one must *develop some actually tractable (or at least nearly tractable) computational technique*—as opposed to deploying euphemisms such as "enabling technology" to obscure a *failure* to devise such a technique.